Deep Learning

CS-898BD

Yassine Ibork

Y723u998

Assignment 2

Supervised by: Dr. Lokesh Das

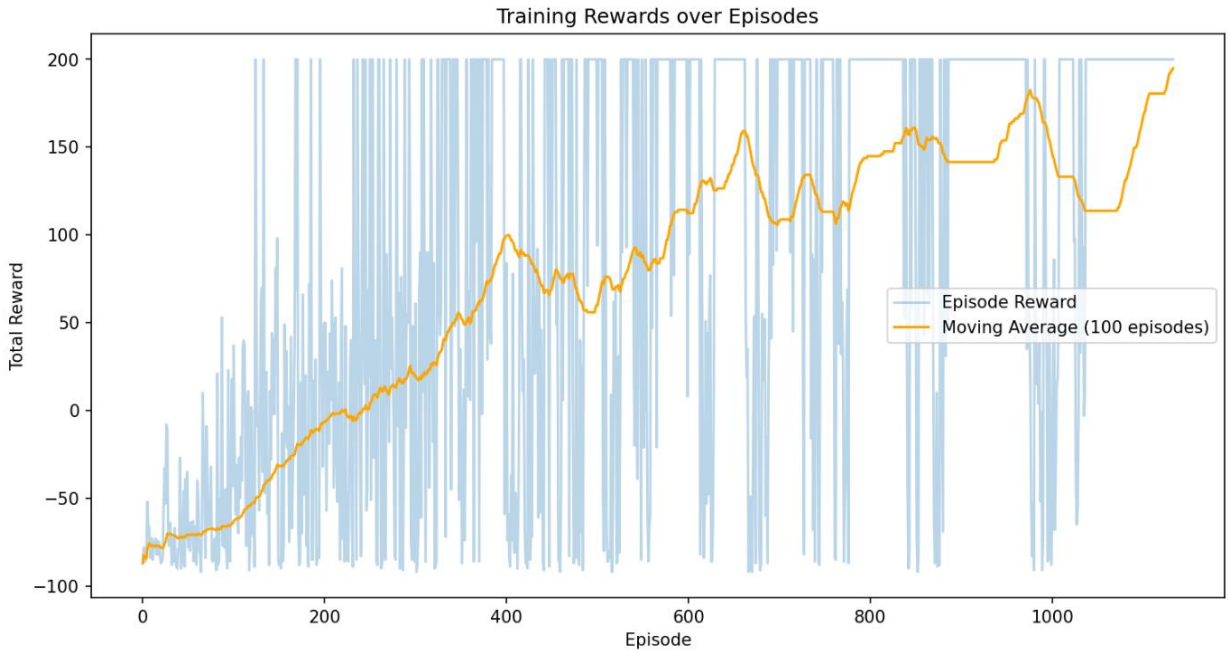
I. Introduction

This assignment help us to get familiar with the reinforcement learning framework by adopting different algorithms on open ai gymnasium rl environment.

The assignment code can be found in my GitHub repository: https://github.com/yibork/DL_RL_openai_gymnasium.

II. Task I. Q- Learning on the Cartpole problem

1. Reward graph



From the figure we can see that the agent is learning after each episode until it achieves a total reward of 200 at the last episode.

2. Q-table shape

Since the number of states is infinite for the cartpole problem I had to divide the states into bins. The observation that I receive from the environment are as follow:

- Cart position
- Cart Velocity
- Pole Angle
- Pole Angular Velocity

To reduce the number of states I have used the following assumptions:

- Since the cart position does not really impact the stability of the pole, I have considered it to have a bin size of 6.
- Since the cart velocity does not really impact the stability of the pole, I have considered it to have a bin size of 6.
- Pole Angle is important for the stability of the pole therefore I have chosen a bin size of 12 for the pole angle.
- Pole Angular Velocity influences how the pole angle is changing therefore I have set a bin size of 12 for it.

With this approach I was able to move from a continuous state space to a discrete state space that has the size of states: 6x6x12x12 = 5184. And since the agent can only take two actions then the size of the q table is 5184x2 = 10,368.

3. Q-table

In this section I will present my Q table for the first five episodes and last five episodes:

3.1. First episode:

state	q_value_action_0	q_value_action_1
2-2-6-6	0.4535	0.564148
2-2-7-6	0	0.1
2-2-7-8	0	0.1
2-2-8-8	-10	0
2-3-6-5	0.56056	0.600798
2-3-6-6	0	0.155495
2-3-7-5	0.13057	0.1
3-2-6-6	0.282762	0
3-2-6-7	0.1	0.11881
3-2-7-7	0.19	0
3-3-5-6	0.1	0
3-3-6-5	0.27991	0

3.2. Second Episode:

state	q_value_action_0	q_value_action_1
2-2-6-6	0.810259	0.756704
2-2-6-7	0.1	0.1
2-2-7-6	0.11088	0.1
2-2-7-7	0.1	0.1099
2-2-7-8	0	0.1
2-2-8-7	0.1099	-9.91
2-2-8-8	-10	0
2-3-6-5	0.679418	0.600798
2-3-6-6	0	0.155495
2-3-7-5	0.13057	0.1
3-2-6-6	0.282762	0
3-2-6-7	0.1	0.11881
3-2-7-7	0.19	0
3-3-5-6	0.1	0
3-3-6-5	0.27991	0

3.3. Third Episode:

state	q_value_action_0	q_value_action_1
2-2-5-5	0	0.1
2-2-6-6	0.810259	0.756704
2-2-6-7	0.1	0.1
2-2-7-6	0.11088	0.1
2-2-7-7	0.1	0.1099
2-2-7-8	0	0.1
2-2-8-7	0.1099	-9.91
2-2-8-8	-10	0
2-3-5-4	0	0.1
2-3-5-5	0	0.19
2-3-6-5	0.679418	0.600798
2-3-6-6	0	0.155495
2-3-7-5	0.13057	0.1
3-2-6-6	0.282762	0
3-2-6-7	0.1	0.11881
3-2-7-7	0.19	0

3-3-3-3	0	-10
3-3-3-4	0	0.19
3-3-4-4	0.19	0.1099
3-3-4-5	0	0.1099
3-3-5-4	0.1	0
3-3-5-6	0.1	0
3-3-6-5	0.27991	0

3.4. Fourth Episode:

state	q_value_action_0	q_value_action_1
2-2-5-5	0	0.1
2-2-6-6	0.810259	0.756704
2-2-6-7	0.1	0.1
2-2-7-6	0.11088	0.1
2-2-7-7	0.1	0.1099
2-2-7-8	0	0.1
2-2-8-7	0.1099	-9.91
2-2-8-8	-10	0
2-3-5-4	0	0.1
2-3-5-5	0	0.19
2-3-6-5	0.679418	0.600798
2-3-6-6	0	0.155495
2-3-7-5	0.13057	0.1
3-2-4-5	0.200977	0.40065
3-2-4-6	0	0.110977
3-2-5-5	0.1	0.404567
3-2-5-6	0.1	0.20881
3-2-6-6	0.282762	0
3-2-6-7	0.1	0.11881
3-2-7-7	0.19	0
3-3-3-3	-10	-10
3-3-3-4	0	0.271
3-3-4-4	0.19	0.21772
3-3-4-5	0.230441	0.21772
3-3-5-4	0.1	0
3-3-5-5	0.573496	0.26135
3-3-5-6	0.1	0
3-3-6-5	0.27991	0

3.5. Fifth Episode

state	q_value_action_0	q_value_action_1
2-2-5-5	0	0.1
2-2-6-6	0.810259	0.756704
2-2-6-7	0.1	0.1
2-2-7-6	0.11088	0.1
2-2-7-7	0.291672	0.218797
2-2-7-8	0	0.1
2-2-8-7	-9.81119	-8.80812
2-2-8-8	-10	0
2-3-5-4	0	0.1
2-3-5-5	0	0.19
2-3-6-5	0.679418	0.600798
2-3-6-6	0	0.155495
2-3-7-5	0.13057	0.1
3-2-4-5	0.200977	0.40065
3-2-4-6	0	0.110977
3-2-5-5	0.1	0.404567
3-2-5-6	0.1	0.20881
3-2-6-6	0.545649	0.137865
3-2-6-7	0.201762	0.217809
3-2-7-7	0.19	0
3-3-3-3	-10	-10
3-3-3-4	0	0.271
3-3-4-4	0.19	0.21772
3-3-4-5	0.230441	0.21772
3-3-5-4	0.1	0
3-3-5-5	0.573496	0.26135
3-3-5-6	0.1	0
3-3-6-5	0.27991	0
3-3-6-6	0.127993	0

3.6. Episode 996

state	q_value_action_0	q_value_action_1
1-1-3-5	0.54998	-17.8237
1-1-3-6	-9.91	0
1-1-4-4	0.129767	0
1-1-4-5	2.391722	0.343914
1-1-4-6	-8.4982	-7.98433
1-1-5-6	0.519303	4.467538
1-1-5-7	0	2.08566
1-1-6-6	1.581646	0.686603
1-1-6-7	-3.07295	-8.31862

1-1-6-8	0	-0.51618	2-1-8-8	-27.1	-18.7821
1-1-7-6	-0.3726	0.122892	2-1-8-9	-19	-10
1-1-7-7	-11.5885	-9.50238	2-1-9-8	-10	0
1-1-7-8	-11.2309	-12.0875	2-2-2-5	-10	0
1-1-8-6	-3.02532	0	2-2-3-4	-50.3045	-46.48
1-1-8-7	-45.1141	-55.0719	2-2-3-5	16.29476	-30.0504
1-1-8-8	-50.442	-52.0706	2-2-3-6	39.32083	21.7508
1-1-8-9	-27.1	-27.1	2-2-3-7	7.09514	20.10113
1-2-2-5	-10	0	2-2-4-4	22.41278	-3.90771
1-2-3-3	-19	0	2-2-4-5	79.89205	64.60373
1-2-3-4	0.308601	-9.57838	2-2-4-6	84.72675	77.01218
1-2-3-5	0.360829	0.136515	2-2-4-7	42.70665	85.42016
1-2-4-4	0.359079	0.1099	2-2-5-4	47.76005	0.321279
1-2-4-5	0.416492	1.121441	2-2-5-5	86.50978	81.00396
1-2-4-6	0.524819	0	2-2-5-6	89.0954	81.38265
1-2-5-5	4.236746	0	2-2-5-7	86.34667	89.02625
1-2-5-6	4.405495	2.739767	2-2-6-4	41.45567	0
1-2-5-7	0.297053	0	2-2-6-5	90.61578	89.90432
1-2-6-5	1.305543	0	2-2-6-6	90.18387	90.58003
1-2-6-6	1.438552	0.589883	2-2-6-7	85.58989	88.19265
1-2-6-7	-0.5208	0.1	2-2-6-8	-0.15556	11.56465
1-2-7-5	0.127711	0	2-2-7-5	73.70293	89.18542
1-2-7-6	-0.21352	0.45431	2-2-7-6	77.32526	88.25819
1-2-7-7	-0.17595	-0.45969	2-2-7-7	53.65336	65.9318
1-2-8-5	1.288501	0	2-2-7-8	-4.91084	-1.85207
1-2-8-6	-3.79937	-8.67672	2-2-8-5	14.5165	70.08433
1-2-8-7	-6.5499	-9.91	2-2-8-6	1.127615	44.10856
1-2-8-8	-10	0	2-2-8-7	-48.0463	24.95629
2-1-3-5	1.704304	0	2-2-8-8	-44.6478	-43.969
2-1-3-6	4.858519	4.514628	2-2-9-7	-10	-10
2-1-3-7	0	2.51912	2-3-2-3	-10	0
2-1-4-5	10.24568	0.694938	2-3-2-4	-10	0
2-1-4-6	8.207735	36.70052	2-3-3-2	0	-10
2-1-4-7	2.065582	51.06669	2-3-3-3	-46.2829	-49.5378
2-1-5-6	5.92148	43.9819	2-3-3-4	-55.911	-63.969
2-1-5-7	1.519114	58.22616	2-3-3-5	1.374351	-4.03136
2-1-5-8	0	0.1	2-3-3-6	3.189758	0
2-1-6-6	-0.01116	0	2-3-4-3	-8.22418	-3.3614
2-1-6-7	0	23.68595	2-3-4-4	43.79733	-10.3237
2-1-6-8	0.787714	0.134046	2-3-4-5	72.07455	28.3537
2-1-7-7	0.19	0	2-3-4-6	33.61846	7.611125
2-1-7-8	-0.58049	-1.7712	2-3-5-3	-0.37516	0.035375
2-1-7-9	0.1	-0.8	2-3-5-4	83.65193	62.09025
2-1-8-7	0.1	0	2-3-5-5	85.78135	83.86026

2-3-5-6	88.13071	65.38418	3-3-2-4	-10	0
2-3-6-3	71.13961	0.853288	3-3-2-5	0	-10
2-3-6-4	89.20899	81.91716	3-3-3-2	-10	0
2-3-6-5	90.56777	90.20627	3-3-3-3	-68.6376	-65.9373
2-3-6-6	90.06171	90.46378	3-3-3-4	25.63284	-60.7927
2-3-7-4	88.26134	80.82342	3-3-3-5	52.0965	5.730588
2-3-7-5	87.1375	89.80215	3-3-3-6	60.49944	1.141368
2-3-7-6	81.39241	88.65964	3-3-4-3	-9.24316	-12.4696
2-3-8-4	72.86191	8.001219	3-3-4-4	80.50339	24.2337
2-3-8-5	57.85479	78.03356	3-3-4-5	88.9594	74.59819
2-3-8-6	21.81664	69.64848	3-3-4-6	89.4669	76.31997
2-3-9-6	0	-10	3-3-4-7	8.319039	0
2-4-3-2	0.19	0	3-3-5-3	19.94755	0.112864
2-4-3-3	-19	-8.9	3-3-5-4	89.65426	81.78931
2-4-4-2	-0.89	0	3-3-5-5	90.30999	89.019
2-4-4-3	-0.05537	0.1	3-3-5-6	89.86215	90.32155
2-4-5-3	0	0.3439	3-3-5-7	7.477421	0
2-4-6-3	25.36711	0	3-3-6-3	28.13741	0.1
2-4-6-4	12.83513	0	3-3-6-4	89.99533	87.2176
2-4-7-4	7.781246	0	3-3-6-5	89.44537	90.0082
3-2-3-4	1.111252	0	3-3-6-6	83.83995	88.67506
3-2-3-5	38.46567	3.830303	3-3-6-7	11.97381	4.616935
3-2-3-6	70.01424	48.83215	3-3-7-4	12.94209	73.47191
3-2-3-7	7.262674	61.93448	3-3-7-5	64.86578	82.47878
3-2-4-4	0	-5.029	3-3-7-6	50.59885	74.08111
3-2-4-5	89.52551	81.8746	3-3-7-7	-2.7479	17.94788
3-2-4-6	89.25762	88.69095	3-3-8-4	16.06664	1.639573
3-2-4-7	74.06044	89.04528	3-3-8-5	3.645457	32.30799
3-2-5-5	89.84079	88.97748	3-3-8-6	-9.37794	14.40131
3-2-5-6	89.274	89.68346	3-3-8-7	-20.1145	-8.06657
3-2-5-7	78.48293	87.77988	3-3-8-8	-10	-19
3-2-6-5	69.61417	89.11004	3-3-9-5	0	0.159274
3-2-6-6	79.65456	88.48531	3-3-9-6	-10	0
3-2-6-7	31.64228	79.07788	3-3-9-7	-10	-10
3-2-6-8	0	8.487369	3-4-2-3	-10	0
3-2-7-5	10.45272	0.136745	3-4-3-2	-10	-10
3-2-7-6	14.87801	63.71243	3-4-3-3	-5.86197	0.1
3-2-7-7	-2.14435	21.28808	3-4-4-2	0.1	0
3-2-7-8	-0.96877	-0.30436	3-4-4-3	-6.24955	0
3-2-8-6	-0.5374	0.610978	3-4-4-4	14.08607	0
3-2-8-7	-20.9033	-11.2663	3-4-4-5	4.016251	0
3-2-8-8	-26.3059	-20.5863	3-4-5-3	-0.57781	0.1
3-2-9-7	-10	0	3-4-5-4	43.42978	0.551769
3-2-9-8	-10	0	3-4-5-5	12.65256	1.266877

3-4-6-3	6.855247	0.1
3-4-6-4	67.93846	6.541029
3-4-6-5	60.7985	4.311784
3-4-6-6	4.058632	0.462662
3-4-7-4	56.90831	4.433232
3-4-7-5	46.54014	23.25323
3-4-7-6	23.11708	4.229402
3-4-8-4	20.76957	0.235868
3-4-8-5	3.539915	16.66392
3-4-8-6	-0.25093	6.927665
4-2-4-6	0.1	0
4-2-4-7	0	0.1
4-2-5-6	0.32289	0
4-2-5-7	0.28803	1.050999
4-2-6-6	0.289877	0
4-2-6-7	0.924122	0.420214
4-2-6-8	0.36942	0
4-2-7-7	-2.23905	-3.48275
4-2-7-8	-2.30144	-2.26819
4-2-7-9	-0.8	0
4-2-8-7	-9.90109	-3.34713
4-2-8-8	-32.594	-31.1322
4-2-8-9	-10	-19
4-3-3-4	0	0.1099
4-3-4-4	0	0.19
4-3-4-5	1.209747	0.215822
4-3-4-6	1.35943	0.130552
4-3-4-7	0.240815	0
4-3-5-4	0.268473	0
4-3-5-5	2.079466	0.380153
4-3-5-6	2.414532	-6.02721
4-3-5-7	1.205873	0
4-3-6-4	0	0.19
4-3-6-5	0.942487	0.308835
4-3-6-6	2.469307	0.602754
4-3-6-7	0.792227	1.381175
4-3-6-8	0.1	0
4-3-7-5	4.356291	0
4-3-7-6	0.486062	0.573763
4-3-7-7	-2.00673	-3.77141
4-3-7-8	0	0.266806
4-3-8-5	-1.17608	0
4-3-8-6	-30.1877	-21.0989
4-3-8-7	-35.6132	-41.3375
4-3-8-8	-17.8051	-28.0687
4-3-9-5	-1.32886	0
4-3-9-7	-10	0
4-4-3-4	-9.91	0
4-4-4-4	0.308885	0
4-4-4-5	-10	0
4-4-4-6	0.1	0
4-4-5-3	-10	0.1
4-4-5-4	0.524693	0
4-4-5-5	0.918745	0.302096
4-4-5-6	0.311211	0
4-4-6-4	-0.25811	0.637956
4-4-6-5	-5.43466	-8.10809
4-4-6-6	0.703622	0
4-4-7-4	0.210356	0.569264
4-4-7-5	-4.83825	-8.94332
4-4-7-6	-1.66593	0.200471
4-4-8-4	0.593303	0
4-4-8-5	-0.23671	1.023826
4-4-8-6	-11.2494	-4.89499
4-5-6-4	0.145399	-10

3.7. Episode 997:

state	q_value_action_0	q_value_action_1
1-1-3-5	0.54998	-17.8237
1-1-3-6	-9.91	0
1-1-4-4	0.129767	0
1-1-4-5	2.391722	0.343914
1-1-4-6	-8.4982	-7.98433
1-1-5-6	0.519303	4.467538
1-1-5-7	0	2.08566
1-1-6-6	1.581646	0.686603
1-1-6-7	-3.07295	-8.31862
1-1-6-8	0	-0.51618
1-1-7-6	-0.3726	0.122892
1-1-7-7	-11.5885	-9.50238
1-1-7-8	-11.2309	-12.0875
1-1-8-6	-3.02532	0
1-1-8-7	-45.1141	-55.0719
1-1-8-8	-50.442	-52.0706
1-1-8-9	-27.1	-27.1
1-2-2-5	-10	0

1-2-3-3	-19	0	2-2-4-5	79.89205	64.60373
1-2-3-4	0.308601	-9.57838	2-2-4-6	84.72675	77.01218
1-2-3-5	0.360829	0.136515	2-2-4-7	42.70665	85.42016
1-2-4-4	0.359079	0.1099	2-2-5-4	47.76005	0.321279
1-2-4-5	0.416492	1.121441	2-2-5-5	86.77925	81.00396
1-2-4-6	0.524819	0	2-2-5-6	89.03824	81.38265
1-2-5-5	4.236746	0	2-2-5-7	86.34667	89.04641
1-2-5-6	4.405495	2.739767	2-2-6-4	41.45567	0
1-2-5-7	0.297053	0	2-2-6-5	90.66688	89.90432
1-2-6-5	1.305543	0	2-2-6-6	90.28844	90.50734
1-2-6-6	1.438552	0.589883	2-2-6-7	85.58989	88.44081
1-2-6-7	-0.5208	0.1	2-2-6-8	-0.15556	11.56465
1-2-7-5	0.127711	0	2-2-7-5	73.70293	89.34282
1-2-7-6	-0.21352	0.45431	2-2-7-6	77.32526	88.25819
1-2-7-7	-0.17595	-0.45969	2-2-7-7	53.65336	65.9318
1-2-8-5	1.288501	0	2-2-7-8	-4.91084	-1.85207
1-2-8-6	-3.79937	-8.67672	2-2-8-5	14.5165	70.08433
1-2-8-7	-6.5499	-9.91	2-2-8-6	1.127615	44.10856
1-2-8-8	-10	0	2-2-8-7	-48.0463	24.95629
2-1-3-5	1.704304	0	2-2-8-8	-44.6478	-43.969
2-1-3-6	4.858519	4.514628	2-2-9-7	-10	-10
2-1-3-7	0	2.51912	2-3-2-3	-10	0
2-1-4-5	10.24568	0.694938	2-3-2-4	-10	0
2-1-4-6	8.207735	36.70052	2-3-3-2	0	-10
2-1-4-7	2.065582	51.06669	2-3-3-3	-46.2829	-49.5378
2-1-5-6	5.92148	43.9819	2-3-3-4	-55.911	-63.969
2-1-5-7	1.519114	58.22616	2-3-3-5	1.374351	-4.03136
2-1-5-8	0	0.1	2-3-3-6	3.189758	0
2-1-6-6	-0.01116	0	2-3-4-3	-8.22418	-3.3614
2-1-6-7	0	23.68595	2-3-4-4	43.79733	-10.3237
2-1-6-8	0.787714	0.134046	2-3-4-5	72.07455	28.3537
2-1-7-7	0.19	0	2-3-4-6	33.61846	7.611125
2-1-7-8	-0.58049	-1.7712	2-3-5-3	-0.37516	0.035375
2-1-7-9	0.1	-0.8	2-3-5-4	83.65193	62.09025
2-1-8-7	0.1	0	2-3-5-5	85.78135	83.86026
2-1-8-8	-27.1	-18.7821	2-3-5-6	88.13071	65.38418
2-1-8-9	-19	-10	2-3-6-3	71.13961	0.853288
2-1-9-8	-10	0	2-3-6-4	89.20899	81.91716
2-2-2-5	-10	0	2-3-6-5	90.67264	90.30984
2-2-3-4	-50.3045	-46.48	2-3-6-6	90.11487	90.54528
2-2-3-5	16.29476	-30.0504	2-3-7-4	88.26134	80.82342
2-2-3-6	39.32083	21.7508	2-3-7-5	87.1375	89.99224
2-2-3-7	7.09514	20.10113	2-3-7-6	81.39241	88.78409
2-2-4-4	22.41278	-3.90771	2-3-8-4	72.86191	8.001219

2-3-8-5	57.85479	78.03356	3-3-4-5	88.9594	74.59819
2-3-8-6	21.81664	69.64848	3-3-4-6	89.4669	76.31997
2-3-9-6	0	-10	3-3-4-7	8.319039	0
2-4-3-2	0.19	0	3-3-5-3	19.94755	0.112864
2-4-3-3	-19	-8.9	3-3-5-4	89.65426	81.78931
2-4-4-2	-0.89	0	3-3-5-5	90.13235	89.28703
2-4-4-3	-0.05537	0.1	3-3-5-6	89.85076	90.47845
2-4-5-3	0	0.3439	3-3-5-7	7.477421	0
2-4-6-3	25.36711	0	3-3-6-3	28.13741	0.1
2-4-6-4	12.83513	0	3-3-6-4	90.06017	87.2176
2-4-7-4	7.781246	0	3-3-6-5	89.44537	90.06294
3-2-3-4	1.111252	0	3-3-6-6	83.83995	88.67506
3-2-3-5	38.46567	3.830303	3-3-6-7	11.97381	4.616935
3-2-3-6	70.01424	48.83215	3-3-7-4	12.94209	73.47191
3-2-3-7	7.262674	61.93448	3-3-7-5	64.86578	82.47878
3-2-4-4	0	-5.029	3-3-7-6	50.59885	74.08111
3-2-4-5	89.50946	81.8746	3-3-7-7	-2.7479	17.94788
3-2-4-6	89.35339	88.85048	3-3-8-4	16.06664	1.639573
3-2-4-7	74.06044	89.04528	3-3-8-5	3.645457	32.30799
3-2-5-5	89.98568	88.97748	3-3-8-6	-9.37794	14.40131
3-2-5-6	89.34359	90.04349	3-3-8-7	-20.1145	-8.06657
3-2-5-7	78.48293	87.77988	3-3-8-8	-10	-19
3-2-6-5	69.61417	89.11004	3-3-9-5	0	0.159274
3-2-6-6	79.65456	88.64759	3-3-9-6	-10	0
3-2-6-7	31.64228	79.07788	3-3-9-7	-10	-10
3-2-6-8	0	8.487369	3-4-2-3	-10	0
3-2-7-5	10.45272	0.136745	3-4-3-2	-10	-10
3-2-7-6	14.87801	63.71243	3-4-3-3	-5.86197	0.1
3-2-7-7	-2.14435	21.28808	3-4-4-2	0.1	0
3-2-7-8	-0.96877	-0.30436	3-4-4-3	-6.24955	0
3-2-8-6	-0.5374	0.610978	3-4-4-4	14.08607	0
3-2-8-7	-20.9033	-11.2663	3-4-4-5	4.016251	0
3-2-8-8	-26.3059	-20.5863	3-4-5-3	-0.57781	0.1
3-2-9-7	-10	0	3-4-5-4	43.42978	0.551769
3-2-9-8	-10	0	3-4-5-5	12.65256	1.266877
3-3-2-4	-10	0	3-4-6-3	6.855247	0.1
3-3-2-5	0	-10	3-4-6-4	67.93846	6.541029
3-3-3-2	-10	0	3-4-6-5	60.7985	4.311784
3-3-3-3	-68.6376	-65.9373	3-4-6-6	4.058632	0.462662
3-3-3-4	25.63284	-60.7927	3-4-7-4	56.90831	4.433232
3-3-3-5	52.0965	5.730588	3-4-7-5	46.54014	23.25323
3-3-3-6	60.49944	1.141368	3-4-7-6	23.11708	4.229402
3-3-4-3	-9.24316	-12.4696	3-4-8-4	20.76957	0.235868
3-3-4-4	80.50339	24.2337	3-4-8-5	3.539915	16.66392

3-4-8-6	-0.25093	6.927665
4-2-4-6	0.1	0
4-2-4-7	0	0.1
4-2-5-6	0.32289	0
4-2-5-7	0.28803	1.050999
4-2-6-6	0.289877	0
4-2-6-7	0.924122	0.420214
4-2-6-8	0.36942	0
4-2-7-7	-2.23905	-3.48275
4-2-7-8	-2.30144	-2.26819
4-2-7-9	-0.8	0
4-2-8-7	-9.90109	-3.34713
4-2-8-8	-32.594	-31.1322
4-2-8-9	-10	-19
4-3-3-4	0	0.1099
4-3-4-4	0	0.19
4-3-4-5	1.209747	0.215822
4-3-4-6	1.35943	0.130552
4-3-4-7	0.240815	0
4-3-5-4	0.268473	0
4-3-5-5	2.079466	0.380153
4-3-5-6	2.414532	-6.02721
4-3-5-7	1.205873	0
4-3-6-4	0	0.19
4-3-6-5	0.942487	0.308835
4-3-6-6	2.469307	0.602754
4-3-6-7	0.792227	1.381175
4-3-6-8	0.1	0
4-3-7-5	4.356291	0
4-3-7-6	0.486062	0.573763
4-3-7-7	-2.00673	-3.77141
4-3-7-8	0	0.266806
4-3-8-5	-1.17608	0
4-3-8-6	-30.1877	-21.0989
4-3-8-7	-35.6132	-41.3375
4-3-8-8	-17.8051	-28.0687
4-3-9-5	-1.32886	0
4-3-9-7	-10	0
4-4-3-4	-9.91	0
4-4-4-4	0.308885	0
4-4-4-5	-10	0
4-4-4-6	0.1	0
4-4-5-3	-10	0.1
4-4-5-4	0.524693	0
4-4-5-5	0.918745	0.302096
4-4-5-6	0.311211	0
4-4-6-4	-0.25811	0.637956
4-4-6-5	-5.43466	-8.10809
4-4-6-6	0.703622	0
4-4-7-4	0.210356	0.569264
4-4-7-5	-4.83825	-8.94332
4-4-7-6	-1.66593	0.200471
4-4-8-4	0.593303	0
4-4-8-5	-0.23671	1.023826
4-4-8-6	-11.2494	-4.89499
4-5-6-4	0.145399	-10

3.8. Episode 998:

state	q_value_action_0	q_value_action_1
1-1-3-5	0.54998	-17.8237
1-1-3-6	-9.91	0
1-1-4-4	0.129767	0
1-1-4-5	2.391722	0.343914
1-1-4-6	-8.4982	-7.98433
1-1-5-6	0.519303	4.467538
1-1-5-7	0	2.08566
1-1-6-6	1.581646	0.686603
1-1-6-7	-3.07295	-8.31862
1-1-6-8	0	-0.51618
1-1-7-6	-0.3726	0.122892
1-1-7-7	-11.5885	-9.50238
1-1-7-8	-11.2309	-12.0875
1-1-8-6	-3.02532	0
1-1-8-7	-45.1141	-55.0719
1-1-8-8	-50.442	-52.0706
1-1-8-9	-27.1	-27.1
1-2-2-5	-10	0
1-2-3-3	-19	0
1-2-3-4	0.308601	-9.57838
1-2-3-5	0.360829	0.136515
1-2-4-4	0.359079	0.1099
1-2-4-5	0.416492	1.121441
1-2-4-6	0.524819	0
1-2-5-5	4.236746	0
1-2-5-6	4.405495	2.739767
1-2-5-7	0.297053	0
1-2-6-5	1.305543	0
1-2-6-6	1.438552	0.589883

1-2-6-7	-0.5208	0.1	2-2-6-8	-0.15556	11.56465
1-2-7-5	0.127711	0	2-2-7-5	73.70293	89.46758
1-2-7-6	-0.21352	0.45431	2-2-7-6	77.32526	88.37731
1-2-7-7	-0.17595	-0.45969	2-2-7-7	53.65336	65.9318
1-2-8-5	1.288501	0	2-2-7-8	-4.91084	-1.85207
1-2-8-6	-3.79937	-8.67672	2-2-8-5	14.5165	70.08433
1-2-8-7	-6.5499	-9.91	2-2-8-6	1.127615	44.10856
1-2-8-8	-10	0	2-2-8-7	-48.0463	24.95629
2-1-3-5	1.704304	0	2-2-8-8	-44.6478	-43.969
2-1-3-6	4.858519	4.514628	2-2-9-7	-10	-10
2-1-3-7	0	2.51912	2-3-2-3	-10	0
2-1-4-5	10.24568	0.694938	2-3-2-4	-10	0
2-1-4-6	8.207735	36.70052	2-3-3-2	0	-10
2-1-4-7	2.065582	51.06669	2-3-3-3	-46.2829	-49.5378
2-1-5-6	5.92148	43.9819	2-3-3-4	-55.911	-63.969
2-1-5-7	1.519114	58.22616	2-3-3-5	1.374351	-4.03136
2-1-5-8	0	0.1	2-3-3-6	3.189758	0
2-1-6-6	-0.01116	0	2-3-4-3	-8.22418	-3.3614
2-1-6-7	0	23.68595	2-3-4-4	43.79733	-10.3237
2-1-6-8	0.787714	0.134046	2-3-4-5	72.07455	28.3537
2-1-7-7	0.19	0	2-3-4-6	33.61846	7.611125
2-1-7-8	-0.58049	-1.7712	2-3-5-3	-0.37516	0.035375
2-1-7-9	0.1	-0.8	2-3-5-4	83.65193	62.09025
2-1-8-7	0.1	0	2-3-5-5	85.78135	83.86026
2-1-8-8	-27.1	-18.7821	2-3-5-6	88.13071	65.38418
2-1-8-9	-19	-10	2-3-6-3	71.13961	0.853288
2-1-9-8	-10	0	2-3-6-4	89.20899	81.91716
2-2-2-5	-10	0	2-3-6-5	90.49538	90.41313
2-2-3-4	-50.3045	-46.48	2-3-6-6	90.17551	90.6902
2-2-3-5	16.29476	-30.0504	2-3-7-4	88.26134	80.82342
2-2-3-6	39.32083	21.7508	2-3-7-5	87.1375	89.99224
2-2-3-7	7.09514	20.10113	2-3-7-6	81.39241	88.78409
2-2-4-4	22.41278	-3.90771	2-3-8-4	72.86191	8.001219
2-2-4-5	79.89205	64.60373	2-3-8-5	57.85479	78.03356
2-2-4-6	84.72675	77.01218	2-3-8-6	21.81664	69.64848
2-2-4-7	42.70665	85.42016	2-3-9-6	0	-10
2-2-5-4	47.76005	0.321279	2-4-3-2	0.19	0
2-2-5-5	87.04683	81.00396	2-4-3-3	-19	-8.9
2-2-5-6	89.28853	81.38265	2-4-4-2	-0.89	0
2-2-5-7	86.34667	89.04641	2-4-4-3	-0.05537	0.1
2-2-6-4	41.45567	0	2-4-5-3	0	0.3439
2-2-6-5	90.68485	89.90432	2-4-6-3	25.36711	0
2-2-6-6	90.33013	90.44341	2-4-6-4	12.83513	0
2-2-6-7	85.58989	88.66939	2-4-7-4	7.781246	0

3-2-3-4	1.111252	0	3-3-6-6	83.83995	88.82637
3-2-3-5	38.46567	3.830303	3-3-6-7	11.97381	4.616935
3-2-3-6	70.01424	48.83215	3-3-7-4	12.94209	73.47191
3-2-3-7	7.262674	61.93448	3-3-7-5	64.86578	82.47878
3-2-4-4	0	-5.029	3-3-7-6	50.59885	74.08111
3-2-4-5	89.5045	81.8746	3-3-7-7	-2.7479	17.94788
3-2-4-6	89.22742	88.85048	3-3-8-4	16.06664	1.639573
3-2-4-7	74.06044	89.04528	3-3-8-5	3.645457	32.30799
3-2-5-5	90.15574	89.41166	3-3-8-6	-9.37794	14.40131
3-2-5-6	89.55869	90.00893	3-3-8-7	-20.1145	-8.06657
3-2-5-7	78.48293	88.04936	3-3-8-8	-10	-19
3-2-6-5	69.61417	89.11004	3-3-9-5	0	0.159274
3-2-6-6	79.65456	88.64759	3-3-9-6	-10	0
3-2-6-7	31.64228	79.07788	3-3-9-7	-10	-10
3-2-6-8	0	8.487369	3-4-2-3	-10	0
3-2-7-5	10.45272	0.136745	3-4-3-2	-10	-10
3-2-7-6	14.87801	63.71243	3-4-3-3	-5.86197	0.1
3-2-7-7	-2.14435	21.28808	3-4-4-2	0.1	0
3-2-7-8	-0.96877	-0.30436	3-4-4-3	-6.24955	0
3-2-8-6	-0.5374	0.610978	3-4-4-4	14.08607	0
3-2-8-7	-20.9033	-11.2663	3-4-4-5	4.016251	0
3-2-8-8	-26.3059	-20.5863	3-4-5-3	-0.57781	0.1
3-2-9-7	-10	0	3-4-5-4	43.42978	0.551769
3-2-9-8	-10	0	3-4-5-5	12.65256	1.266877
3-3-2-4	-10	0	3-4-6-3	6.855247	0.1
3-3-2-5	0	-10	3-4-6-4	67.93846	6.541029
3-3-3-2	-10	0	3-4-6-5	60.7985	4.311784
3-3-3-3	-68.6376	-65.9373	3-4-6-6	4.058632	0.462662
3-3-3-4	25.63284	-60.7927	3-4-7-4	56.90831	4.433232
3-3-3-5	52.0965	5.730588	3-4-7-5	46.54014	23.25323
3-3-3-6	60.49944	1.141368	3-4-7-6	23.11708	4.229402
3-3-4-3	-9.24316	-12.4696	3-4-8-4	20.76957	0.235868
3-3-4-4	80.50339	24.2337	3-4-8-5	3.539915	16.66392
3-3-4-5	88.9594	74.59819	3-4-8-6	-0.25093	6.927665
3-3-4-6	89.4669	76.31997	4-2-4-6	0.1	0
3-3-4-7	8.319039	0	4-2-4-7	0	0.1
3-3-5-3	19.94755	0.112864	4-2-5-6	0.32289	0
3-3-5-4	89.65426	81.78931	4-2-5-7	0.28803	1.050999
3-3-5-5	90.37136	89.3859	4-2-6-6	0.289877	0
3-3-5-6	89.85076	90.54079	4-2-6-7	0.924122	0.420214
3-3-5-7	7.477421	0	4-2-6-8	0.36942	0
3-3-6-3	28.13741	0.1	4-2-7-7	-2.23905	-3.48275
3-3-6-4	90.09296	87.2176	4-2-7-8	-2.30144	-2.26819
3-3-6-5	89.44537	90.11396	4-2-7-9	-0.8	0

4-2-8-7	-9.90109	-3.34713
4-2-8-8	-32.594	-31.1322
4-2-8-9	-10	-19
4-3-3-4	0	0.1099
4-3-4-4	0	0.19
4-3-4-5	1.209747	0.215822
4-3-4-6	1.35943	0.130552
4-3-4-7	0.240815	0
4-3-5-4	0.268473	0
4-3-5-5	2.079466	0.380153
4-3-5-6	2.414532	-6.02721
4-3-5-7	1.205873	0
4-3-6-4	0	0.19
4-3-6-5	0.942487	0.308835
4-3-6-6	2.469307	0.602754
4-3-6-7	0.792227	1.381175
4-3-6-8	0.1	0
4-3-7-5	4.356291	0
4-3-7-6	0.486062	0.573763
4-3-7-7	-2.00673	-3.77141
4-3-7-8	0	0.266806
4-3-8-5	-1.17608	0
4-3-8-6	-30.1877	-21.0989
4-3-8-7	-35.6132	-41.3375
4-3-8-8	-17.8051	-28.0687
4-3-9-5	-1.32886	0
4-3-9-7	-10	0
4-4-3-4	-9.91	0
4-4-4-4	0.308885	0
4-4-4-5	-10	0
4-4-4-6	0.1	0
4-4-5-3	-10	0.1
4-4-5-4	0.524693	0
4-4-5-5	0.918745	0.302096
4-4-5-6	0.311211	0
4-4-6-4	-0.25811	0.637956
4-4-6-5	-5.43466	-8.10809
4-4-6-6	0.703622	0
4-4-7-4	0.210356	0.569264
4-4-7-5	-4.83825	-8.94332
4-4-7-6	-1.66593	0.200471
4-4-8-4	0.593303	0
4-4-8-5	-0.23671	1.023826
4-4-8-6	-11.2494	-4.89499
4-5-6-4	0.145399	-10

3.9. Episode 999

state	q_value_action_0	q_value_action_1
1-1-3-5	0.54998	-17.8237
1-1-3-6	-9.91	0
1-1-4-4	0.129767	0
1-1-4-5	2.391722	0.343914
1-1-4-6	-8.4982	-7.98433
1-1-5-6	0.519303	4.467538
1-1-5-7	0	2.08566
1-1-6-6	1.581646	0.686603
1-1-6-7	-3.07295	-8.31862
1-1-6-8	0	-0.51618
1-1-7-6	-0.3726	0.122892
1-1-7-7	-11.5885	-9.50238
1-1-7-8	-11.2309	-12.0875
1-1-8-6	-3.02532	0
1-1-8-7	-45.1141	-55.0719
1-1-8-8	-50.442	-52.0706
1-1-8-9	-27.1	-27.1
1-2-2-5	-10	0
1-2-3-3	-19	0
1-2-3-4	0.308601	-9.57838
1-2-3-5	0.360829	0.136515
1-2-4-4	0.359079	0.1099
1-2-4-5	0.416492	1.121441
1-2-4-6	0.524819	0
1-2-5-5	4.236746	0
1-2-5-6	4.405495	2.739767
1-2-5-7	0.297053	0
1-2-6-5	1.305543	0
1-2-6-6	1.438552	0.589883
1-2-6-7	-0.5208	0.1
1-2-7-5	0.127711	0
1-2-7-6	-0.21352	0.45431
1-2-7-7	-0.17595	-0.45969
1-2-8-5	1.288501	0
1-2-8-6	-3.79937	-8.67672
1-2-8-7	-6.5499	-9.91
1-2-8-8	-10	0
2-1-3-5	1.704304	0
2-1-3-6	4.858519	4.514628
2-1-3-7	0	2.51912

2-1-4-5	10.24568	0.694938	2-3-2-4	-10	0
2-1-4-6	8.207735	36.70052	2-3-3-2	0	-10
2-1-4-7	2.065582	51.06669	2-3-3-3	-46.2829	-49.5378
2-1-5-6	5.92148	43.9819	2-3-3-4	-55.911	-63.969
2-1-5-7	1.519114	58.22616	2-3-3-5	1.374351	-4.03136
2-1-5-8	0	0.1	2-3-3-6	3.189758	0
2-1-6-6	-0.01116	0	2-3-4-3	-8.22418	-3.3614
2-1-6-7	0	23.68595	2-3-4-4	43.79733	-10.3237
2-1-6-8	0.787714	0.134046	2-3-4-5	72.07455	28.3537
2-1-7-7	0.19	0	2-3-4-6	33.61846	7.611125
2-1-7-8	-0.58049	-1.7712	2-3-5-3	-0.37516	0.035375
2-1-7-9	0.1	-0.8	2-3-5-4	83.65193	62.09025
2-1-8-7	0.1	0	2-3-5-5	85.78135	83.86026
2-1-8-8	-27.1	-18.7821	2-3-5-6	88.13071	65.38418
2-1-8-9	-19	-10	2-3-6-3	71.13961	0.853288
2-1-9-8	-10	0	2-3-6-4	89.20899	81.91716
2-2-2-5	-10	0	2-3-6-5	90.63576	90.54359
2-2-3-4	-50.3045	-46.48	2-3-6-6	90.17551	90.82385
2-2-3-5	16.29476	-30.0504	2-3-7-4	88.26134	80.82342
2-2-3-6	39.32083	21.7508	2-3-7-5	87.1375	89.99224
2-2-3-7	7.09514	20.10113	2-3-7-6	81.39241	88.78409
2-2-4-4	22.41278	-3.90771	2-3-8-4	72.86191	8.001219
2-2-4-5	79.89205	64.60373	2-3-8-5	57.85479	78.03356
2-2-4-6	84.72675	77.01218	2-3-8-6	21.81664	69.64848
2-2-4-7	42.70665	85.42016	2-3-9-6	0	-10
2-2-5-4	47.76005	0.321279	2-4-3-2	0.19	0
2-2-5-5	87.49311	81.00396	2-4-3-3	-19	-8.9
2-2-5-6	89.24335	81.98527	2-4-4-2	-0.89	0
2-2-5-7	86.34667	89.08193	2-4-4-3	-0.05537	0.1
2-2-6-4	41.45567	0	2-4-5-3	0	0.3439
2-2-6-5	90.67022	89.97293	2-4-6-3	25.36711	0
2-2-6-6	90.35603	90.64098	2-4-6-4	12.83513	0
2-2-6-7	85.58989	88.85635	2-4-7-4	7.781246	0
2-2-6-8	-0.15556	11.56465	3-2-3-4	1.111252	0
2-2-7-5	73.70293	89.46758	3-2-3-5	38.46567	3.830303
2-2-7-6	77.32526	88.37731	3-2-3-6	70.01424	48.83215
2-2-7-7	53.65336	65.9318	3-2-3-7	7.262674	61.93448
2-2-7-8	-4.91084	-1.85207	3-2-4-4	0	-5.029
2-2-8-5	14.5165	70.08433	3-2-4-5	89.5045	81.8746
2-2-8-6	1.127615	44.10856	3-2-4-6	89.23954	88.85048
2-2-8-7	-48.0463	24.95629	3-2-4-7	74.06044	89.15004
2-2-8-8	-44.6478	-43.969	3-2-5-5	90.17635	89.41166
2-2-9-7	-10	-10	3-2-5-6	89.58635	90.15684
2-3-2-3	-10	0	3-2-5-7	78.48293	88.04936

3-2-6-5	69.61417	89.11004	3-3-9-5	0	0.159274
3-2-6-6	79.65456	88.8395	3-3-9-6	-10	0
3-2-6-7	31.64228	79.07788	3-3-9-7	-10	-10
3-2-6-8	0	8.487369	3-4-2-3	-10	0
3-2-7-5	10.45272	0.136745	3-4-3-2	-10	-10
3-2-7-6	14.87801	63.71243	3-4-3-3	-5.86197	0.1
3-2-7-7	-2.14435	21.28808	3-4-4-2	0.1	0
3-2-7-8	-0.96877	-0.30436	3-4-4-3	-6.24955	0
3-2-8-6	-0.5374	0.610978	3-4-4-4	14.08607	0
3-2-8-7	-20.9033	-11.2663	3-4-4-5	4.016251	0
3-2-8-8	-26.3059	-20.5863	3-4-5-3	-0.57781	0.1
3-2-9-7	-10	0	3-4-5-4	43.42978	0.551769
3-2-9-8	-10	0	3-4-5-5	12.65256	1.266877
3-3-2-4	-10	0	3-4-6-3	6.855247	0.1
3-3-2-5	0	-10	3-4-6-4	67.93846	6.541029
3-3-3-2	-10	0	3-4-6-5	60.7985	4.311784
3-3-3-3	-68.6376	-65.9373	3-4-6-6	4.058632	0.462662
3-3-3-4	25.63284	-60.7927	3-4-7-4	56.90831	4.433232
3-3-3-5	52.0965	5.730588	3-4-7-5	46.54014	23.25323
3-3-3-6	60.49944	1.141368	3-4-7-6	23.11708	4.229402
3-3-4-3	-9.24316	-12.4696	3-4-8-4	20.76957	0.235868
3-3-4-4	80.50339	24.2337	3-4-8-5	3.539915	16.66392
3-3-4-5	89.04054	74.59819	3-4-8-6	-0.25093	6.927665
3-3-4-6	89.45372	76.31997	4-2-4-6	0.1	0
3-3-4-7	8.319039	0	4-2-4-7	0	0.1
3-3-5-3	19.94755	0.112864	4-2-5-6	0.32289	0
3-3-5-4	89.79746	81.78931	4-2-5-7	0.28803	1.050999
3-3-5-5	90.59532	89.55568	4-2-6-6	0.289877	0
3-3-5-6	89.92665	90.6209	4-2-6-7	0.924122	0.420214
3-3-5-7	7.477421	0	4-2-6-8	0.36942	0
3-3-6-3	28.13741	0.1	4-2-7-7	-2.23905	-3.48275
3-3-6-4	90.12332	87.2176	4-2-7-8	-2.30144	-2.26819
3-3-6-5	89.44537	90.23594	4-2-7-9	-0.8	0
3-3-6-6	83.83995	89.09832	4-2-8-7	-9.90109	-3.34713
3-3-6-7	11.97381	4.616935	4-2-8-8	-32.594	-31.1322
3-3-7-4	12.94209	73.47191	4-2-8-9	-10	-19
3-3-7-5	64.86578	82.47878	4-3-3-4	0	0.1099
3-3-7-6	50.59885	74.08111	4-3-4-4	0	0.19
3-3-7-7	-2.7479	17.94788	4-3-4-5	1.209747	0.215822
3-3-8-4	16.06664	1.639573	4-3-4-6	1.35943	0.130552
3-3-8-5	3.645457	32.30799	4-3-4-7	0.240815	0
3-3-8-6	-9.37794	14.40131	4-3-5-4	0.268473	0
3-3-8-7	-20.1145	-8.06657	4-3-5-5	2.079466	0.380153
3-3-8-8	-10	-19	4-3-5-6	2.414532	-6.02721

4-3-5-7	1.205873	0
4-3-6-4	0	0.19
4-3-6-5	0.942487	0.308835
4-3-6-6	2.469307	0.602754
4-3-6-7	0.792227	1.381175
4-3-6-8	0.1	0
4-3-7-5	4.356291	0
4-3-7-6	0.486062	0.573763
4-3-7-7	-2.00673	-3.77141
4-3-7-8	0	0.266806
4-3-8-5	-1.17608	0
4-3-8-6	-30.1877	-21.0989
4-3-8-7	-35.6132	-41.3375
4-3-8-8	-17.8051	-28.0687
4-3-9-5	-1.32886	0
4-3-9-7	-10	0
4-4-3-4	-9.91	0
4-4-4-4	0.308885	0
4-4-4-5	-10	0
4-4-4-6	0.1	0
4-4-5-3	-10	0.1
4-4-5-4	0.524693	0
4-4-5-5	0.918745	0.302096
4-4-5-6	0.311211	0
4-4-6-4	-0.25811	0.637956
4-4-6-5	-5.43466	-8.10809
4-4-6-6	0.703622	0
4-4-7-4	0.210356	0.569264
4-4-7-5	-4.83825	-8.94332
4-4-7-6	-1.66593	0.200471
4-4-8-4	0.593303	0
4-4-8-5	-0.23671	1.023826
4-4-8-6	-11.2494	-4.89499
4-5-6-4	0.145399	-10

3.10. Episode 1000

state	q_value_action_0	q_value_action_1
1-1-3-5	0.54998	-17.8237
1-1-3-6	-9.91	0
1-1-4-4	0.129767	0
1-1-4-5	2.391722	0.343914
1-1-4-6	-8.4982	-7.98433
1-1-5-6	0.519303	4.467538
1-1-5-7	0	2.08566
1-1-6-6	1.581646	0.686603
1-1-6-7	-3.07295	-8.31862
1-1-6-8	0	-0.51618
1-1-7-6	-0.3726	0.122892
1-1-7-7	-11.5885	-9.50238
1-1-7-8	-11.2309	-12.0875
1-1-8-6	-3.02532	0
1-1-8-7	-45.1141	-55.0719
1-1-8-8	-50.442	-52.0706
1-1-8-9	-27.1	-27.1
1-2-2-5	-10	0
1-2-3-3	-19	0
1-2-3-4	0.308601	-9.57838
1-2-3-5	0.360829	0.136515
1-2-4-4	0.359079	0.1099
1-2-4-5	0.416492	1.121441
1-2-4-6	0.524819	0
1-2-5-5	4.236746	0
1-2-5-6	4.405495	2.739767
1-2-5-7	0.297053	0
1-2-6-5	1.305543	0
1-2-6-6	1.438552	0.589883
1-2-6-7	-0.5208	0.1
1-2-7-5	0.127711	0
1-2-7-6	-0.21352	0.45431
1-2-7-7	-0.17595	-0.45969
1-2-8-5	1.288501	0
1-2-8-6	-3.79937	-8.67672
1-2-8-7	-6.5499	-9.91
1-2-8-8	-10	0
2-1-3-5	1.704304	0
2-1-3-6	4.858519	4.514628
2-1-3-7	0	2.51912
2-1-4-5	10.24568	0.694938
2-1-4-6	8.207735	36.70052
2-1-4-7	2.065582	51.06669
2-1-5-6	5.92148	43.9819
2-1-5-7	1.519114	58.22616
2-1-5-8	0	0.1
2-1-6-6	-0.01116	0
2-1-6-7	0	23.68595
2-1-6-8	0.787714	0.134046
2-1-7-7	0.19	0
2-1-7-8	-0.58049	-1.7712

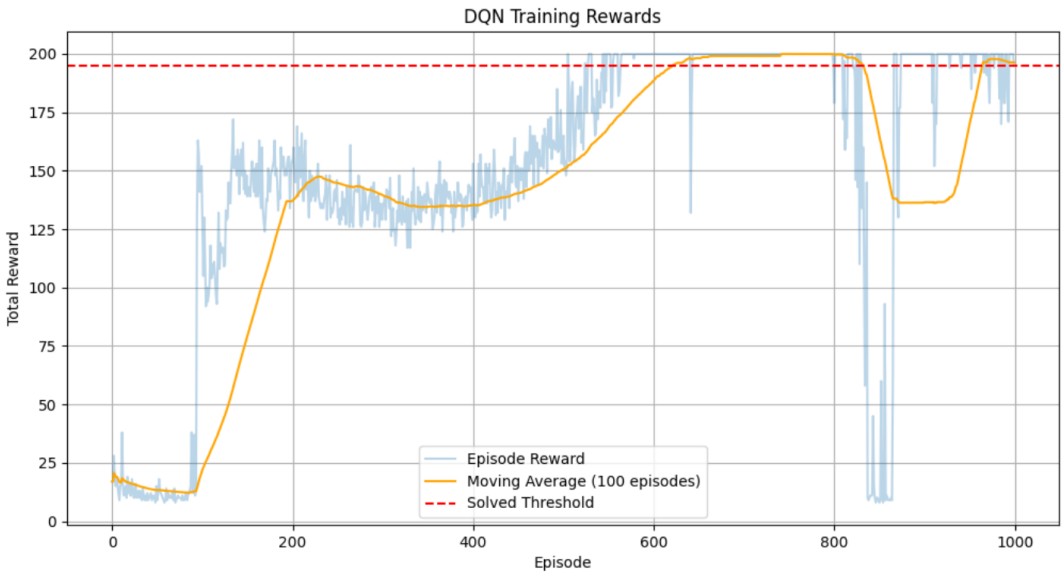
2-1-7-9	0.1	-0.8	2-3-5-4	79.73739	62.09025
2-1-8-7	0.1	0	2-3-5-5	86.13831	83.86026
2-1-8-8	-27.1	-18.7821	2-3-5-6	88.13071	65.38418
2-1-8-9	-19	-10	2-3-6-3	71.13961	0.853288
2-1-9-8	-10	0	2-3-6-4	89.37584	82.78783
2-2-2-5	-10	0	2-3-6-5	90.46663	90.48198
2-2-3-4	-50.3045	-46.48	2-3-6-6	90.29211	90.76612
2-2-3-5	18.65804	-30.0504	2-3-7-4	88.26134	80.82342
2-2-3-6	44.05451	21.7508	2-3-7-5	87.1375	89.9247
2-2-3-7	7.09514	20.10113	2-3-7-6	81.39241	88.78409
2-2-4-4	22.41278	-3.90771	2-3-8-4	72.86191	8.001219
2-2-4-5	79.89205	64.60373	2-3-8-5	57.85479	78.03356
2-2-4-6	84.81906	77.01218	2-3-8-6	21.81664	69.64848
2-2-4-7	42.70665	85.78405	2-3-9-6	0	-10
2-2-5-4	47.76005	0.321279	2-4-3-2	0.19	0
2-2-5-5	87.6823	81.00396	2-4-3-3	-19	-8.9
2-2-5-6	89.42669	81.98527	2-4-4-2	-0.89	0
2-2-5-7	86.60878	89.43621	2-4-4-3	-0.05537	0.1
2-2-6-4	41.45567	0	2-4-5-3	0	0.3439
2-2-6-5	90.81734	90.06317	2-4-6-3	25.36711	0
2-2-6-6	90.43677	90.83396	2-4-6-4	12.83513	0
2-2-6-7	85.58989	89.0665	2-4-7-4	7.781246	0
2-2-6-8	-0.15556	11.56465	3-2-3-4	1.111252	0
2-2-7-5	73.70293	89.53954	3-2-3-5	38.46567	3.830303
2-2-7-6	77.32526	88.50733	3-2-3-6	70.01424	48.83215
2-2-7-7	53.65336	65.9318	3-2-3-7	7.262674	61.93448
2-2-7-8	-4.91084	-1.85207	3-2-4-4	0	-5.029
2-2-8-5	14.5165	70.08433	3-2-4-5	89.48909	81.8746
2-2-8-6	1.127615	44.10856	3-2-4-6	89.51265	88.85048
2-2-8-7	-48.0463	24.95629	3-2-4-7	74.06044	89.15004
2-2-8-8	-44.6478	-43.969	3-2-5-5	90.21685	89.52435
2-2-9-7	-10	-10	3-2-5-6	89.67182	90.36096
2-3-2-3	-10	0	3-2-5-7	78.48293	88.04936
2-3-2-4	-10	0	3-2-6-5	69.61417	89.11004
2-3-3-2	0	-10	3-2-6-6	79.65456	88.8395
2-3-3-3	-46.2829	-49.5378	3-2-6-7	31.64228	79.07788
2-3-3-4	-55.911	-63.969	3-2-6-8	0	8.487369
2-3-3-5	2.950097	-4.03136	3-2-7-5	10.45272	0.136745
2-3-3-6	3.189758	0	3-2-7-6	14.87801	63.71243
2-3-4-3	-8.22418	-3.3614	3-2-7-7	-2.14435	21.28808
2-3-4-4	46.65298	-10.3237	3-2-7-8	-0.96877	-0.30436
2-3-4-5	65.15339	32.75924	3-2-8-6	-0.5374	0.610978
2-3-4-6	33.61846	7.611125	3-2-8-7	-20.9033	-11.2663
2-3-5-3	-0.37516	0.035375	3-2-8-8	-26.3059	-20.5863

3-2-9-7	-10	0	3-4-5-4	43.42978	0.551769
3-2-9-8	-10	0	3-4-5-5	12.65256	1.266877
3-3-2-4	-10	0	3-4-6-3	6.855247	0.1
3-3-2-5	0	-10	3-4-6-4	67.93846	6.541029
3-3-3-2	-10	0	3-4-6-5	60.7985	4.311784
3-3-3-3	-68.6376	-65.9373	3-4-6-6	4.058632	0.462662
3-3-3-4	25.63284	-60.7927	3-4-7-4	56.90831	4.433232
3-3-3-5	52.0965	5.730588	3-4-7-5	46.54014	23.25323
3-3-3-6	60.49944	1.141368	3-4-7-6	23.11708	4.229402
3-3-4-3	-9.24316	-12.4696	3-4-8-4	20.76957	0.235868
3-3-4-4	80.50339	24.2337	3-4-8-5	3.539915	16.66392
3-3-4-5	89.09744	74.59819	3-4-8-6	-0.25093	6.927665
3-3-4-6	89.45372	76.31997	4-2-4-6	0.1	0
3-3-4-7	8.319039	0	4-2-4-7	0	0.1
3-3-5-3	19.94755	0.112864	4-2-5-6	0.32289	0
3-3-5-4	89.79746	81.78931	4-2-5-7	0.28803	1.050999
3-3-5-5	90.43393	89.77266	4-2-6-6	0.289877	0
3-3-5-6	89.92665	90.68468	4-2-6-7	0.924122	0.420214
3-3-5-7	7.477421	0	4-2-6-8	0.36942	0
3-3-6-3	28.13741	0.1	4-2-7-7	-2.23905	-3.48275
3-3-6-4	90.17992	87.2176	4-2-7-8	-2.30144	-2.26819
3-3-6-5	89.44537	90.25212	4-2-7-9	-0.8	0
3-3-6-6	83.83995	89.09832	4-2-8-7	-9.90109	-3.34713
3-3-6-7	11.97381	4.616935	4-2-8-8	-32.594	-31.1322
3-3-7-4	12.94209	73.47191	4-2-8-9	-10	-19
3-3-7-5	64.86578	82.47878	4-3-3-4	0	0.1099
3-3-7-6	50.59885	74.08111	4-3-4-4	0	0.19
3-3-7-7	-2.7479	17.94788	4-3-4-5	1.209747	0.215822
3-3-8-4	16.06664	1.639573	4-3-4-6	1.35943	0.130552
3-3-8-5	3.645457	32.30799	4-3-4-7	0.240815	0
3-3-8-6	-9.37794	14.40131	4-3-5-4	0.268473	0
3-3-8-7	-20.1145	-8.06657	4-3-5-5	2.079466	0.380153
3-3-8-8	-10	-19	4-3-5-6	2.414532	-6.02721
3-3-9-5	0	0.159274	4-3-5-7	1.205873	0
3-3-9-6	-10	0	4-3-6-4	0	0.19
3-3-9-7	-10	-10	4-3-6-5	0.942487	0.308835
3-4-2-3	-10	0	4-3-6-6	2.469307	0.602754
3-4-3-2	-10	-10	4-3-6-7	0.792227	1.381175
3-4-3-3	-5.86197	0.1	4-3-6-8	0.1	0
3-4-4-2	0.1	0	4-3-7-5	4.356291	0
3-4-4-3	-6.24955	0	4-3-7-6	0.486062	0.573763
3-4-4-4	14.08607	0	4-3-7-7	-2.00673	-3.77141
3-4-4-5	4.016251	0	4-3-7-8	0	0.266806
3-4-5-3	-0.57781	0.1	4-3-8-5	-1.17608	0

4-3-8-6	-30.1877	-21.0989	4-4-5-6	0.311211	0
4-3-8-7	-35.6132	-41.3375	4-4-6-4	-0.25811	0.637956
4-3-8-8	-17.8051	-28.0687	4-4-6-5	-5.43466	-8.10809
4-3-9-5	-1.32886	0	4-4-6-6	0.703622	0
4-3-9-7	-10	0	4-4-7-4	0.210356	0.569264
4-4-3-4	-9.91	0	4-4-7-5	-4.83825	-8.94332
4-4-4-4	0.308885	0	4-4-7-6	-1.66593	0.200471
4-4-4-5	-10	0	4-4-8-4	0.593303	0
4-4-4-6	0.1	0	4-4-8-5	-0.23671	1.023826
4-4-5-3	-10	0.1	4-4-8-6	-11.2494	-4.89499
4-4-5-4	0.524693	0	4-5-6-4	0.145399	-10
4-4-5-5	0.918745	0.302096			

III. Task II. Deep Q- Learning on the Cartpole problem

1. Reward Graph



From the figure we can see that the agent is learning an episode after the other it reached the maximum reward at episode 600 so we could have stopped learning in that episode to not loose computation power.

2. Model Architecture

```python
class DQNNetwork(nn.Module):
    """
    Deep Q-Network (DQN) Architecture.

    Parameters:
    - state_dim (int): Dimension of the state space.
    - action_dim (int): Number of possible actions.
    - hidden_dim (int): Number of neurons in hidden layers.
    """
    def __init__(self, state_dim, action_dim, hidden_dim=128):
        super(DQNNetwork, self).__init__()
        self.fc1 = nn.Linear(state_dim, hidden_dim)
        self.fc2 = nn.Linear(hidden_dim, hidden_dim)
        self.out = nn.Linear(hidden_dim, action_dim)

    def forward(self, x):
        """
        Forward pass through the network.

        Parameters:
        - x (torch.Tensor): Input state tensor.

        Returns:
        - torch.Tensor: Q-values for each action.
        """
        x = F.relu(self.fc1(x))
        x = F.relu(self.fc2(x))
        return self.out(x)
```

The model architecture of that we have used to solve the cartpole problem using deep reinforcement learning is as follows:

Two fully connected layers:

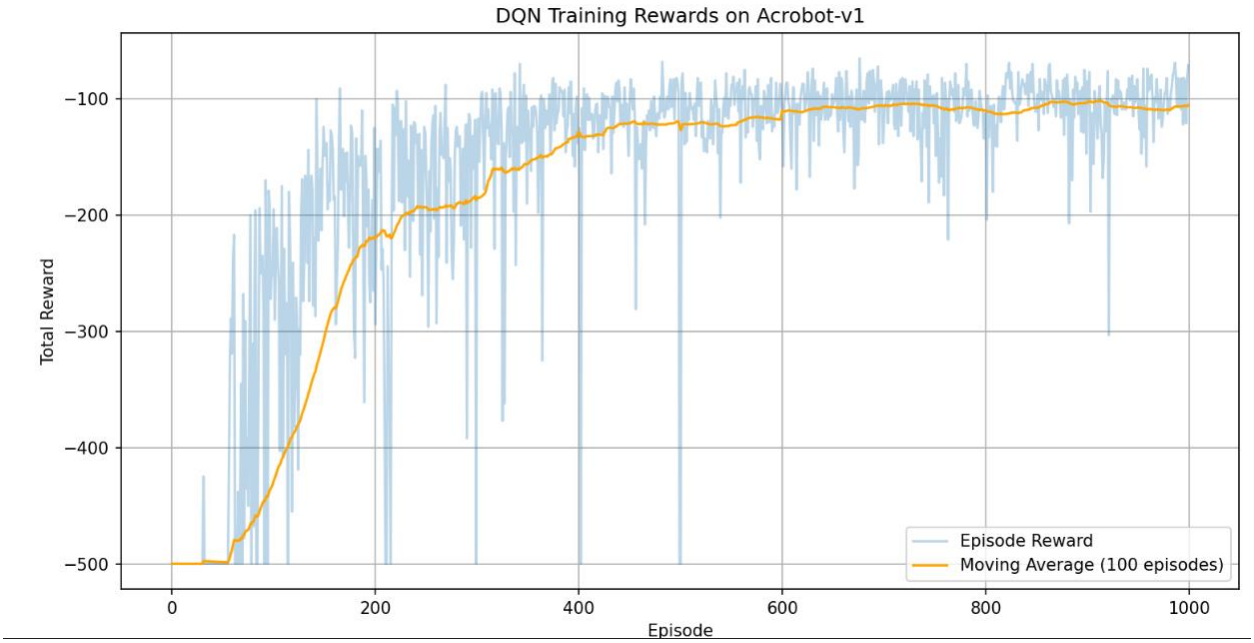
- Fc1: has dimension state dim x 128
- Fc2: has dimension state 128 x 128

One output layer:

- Out: has dimension of 128 x number of actions

IV. Task 3: Deep Q Learning on Acrobot:

1. Reward Graph



Graph reward of the acrobot-v1 shows that the agent learns episode after episode.

2. Model Architecture

```python
class DQNNetwork(nn.Module):
    """
    Deep Q-Network (DQN) Architecture.
    """
    def __init__(self, state_dim, action_dim, hidden_dim=128):
        super(DQNNetwork, self).__init__()
        self.fc1 = nn.Linear(state_dim, hidden_dim)
        self.fc2 = nn.Linear(hidden_dim, hidden_dim)
        self.out = nn.Linear(hidden_dim, action_dim)

    def forward(self, x):
        """
        Forward pass through the network.
        """
        x = F.relu(self.fc1(x))
        x = F.relu(self.fc2(x))
        return self.out(x)
```

The model architecture of that we have used to solve the Acrobot using deep reinforcement learning is as follows:

Two fully connected layers:

- Fc1: has dimension state dim x 128
- Fc2: has dimension state 128 x 128

One output layer:

- Out: has dimension of 128 x number of actions

V. References

https://pytorch.org/tutorials/intermediate/reinforcement_q_learning.html
https://www.gymlibrary.dev/