

Report Project 1

Navigation Project

Tri. B. Minh

15th May, 2020

REPORT PROJECT 1 NAVIGATION ENVIRONMENT

Description of the Banana Navigation Environment

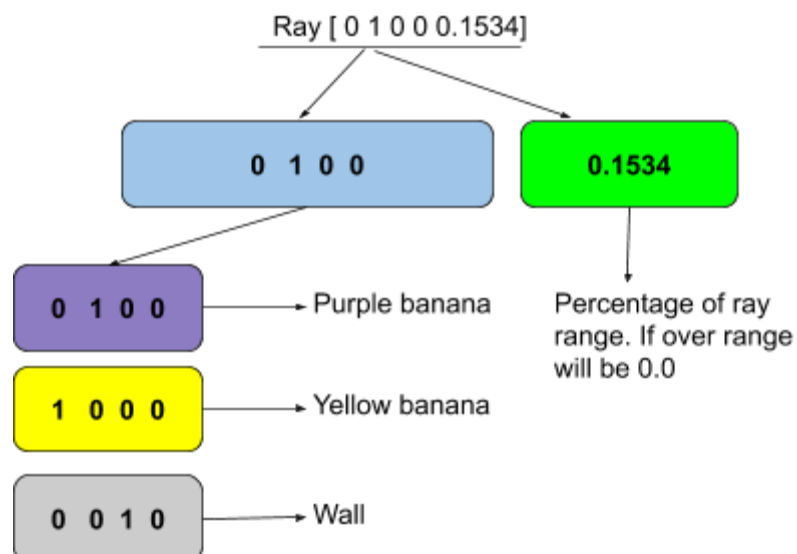
Number of Agent: 1 and moving the planar environment to collect the banana.

Observation: A set of 37 measurement ranges and kinds of objects in the environment.

Action: 4 action [0, 1, 2, 3] with [Move forward, Move Backward, Turn Left, Turn Right]

Reward: +1 when Agent get Yellow Banana and -1 when get Purple Banana

Detail in Observations: 7 ray perceptions with each ray include 5 entries (35 values) and 2 Value agent's velocity



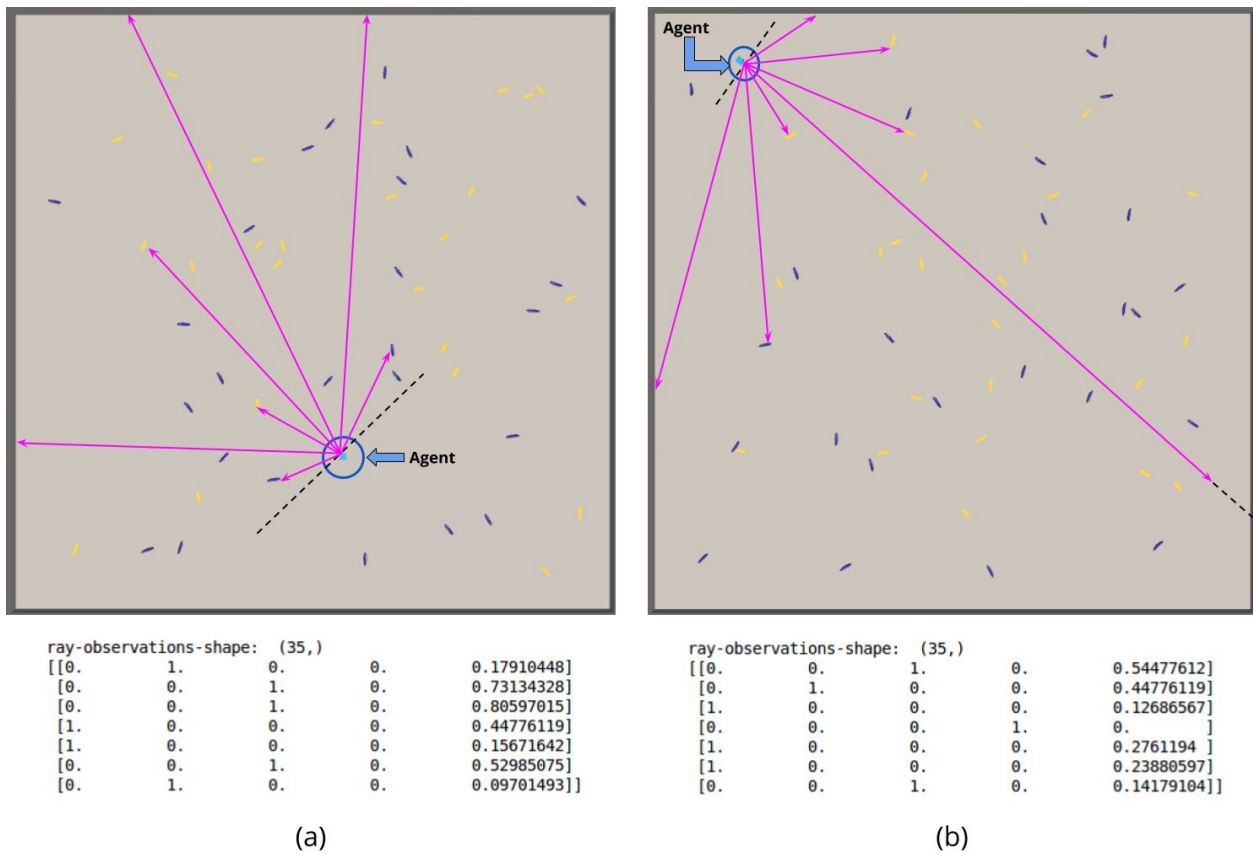
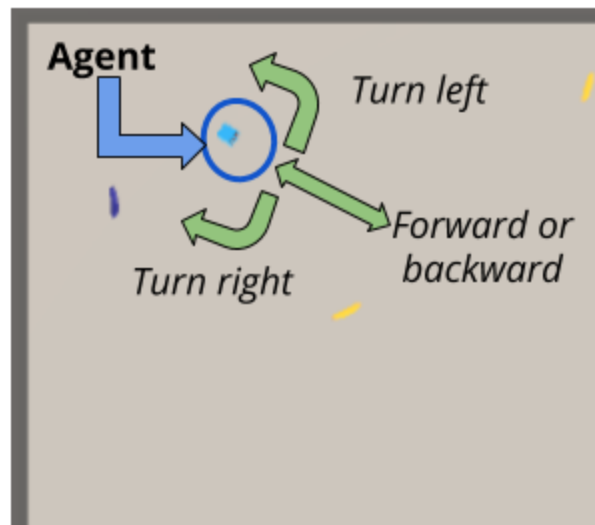


Figure 2. Agent ray-perceptions. a) 7 rays reaching at least one object (banana or wall). b) One ray reaching the max. length before reaching any object ([Source](#))

Detail in Actions:

- **Action 0:** Move forward.
- **Action 1:** Move backward.
- **Action 2:** Turn left.
- **Action 3:** Turn right.



Result: The best result is an environment solved in 529 episodes and get the average score 13.0. All parameters will list below.

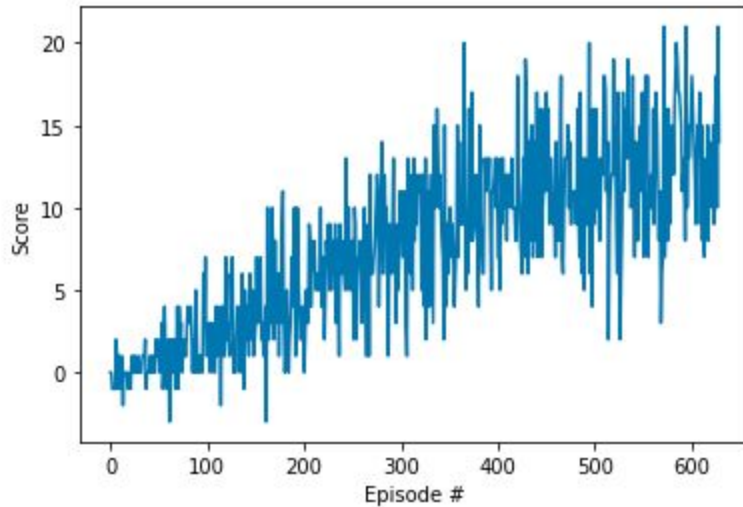
Experiment 1. Natural DQN

<i>n_episode</i>	<i>max_t</i>	<i>eps_start</i>	<i>eps_end</i>	<i>eps_decay</i>	<i>Buffer_Size</i>	<i>Batch Size</i>
2000	1000	1.0	0.01	0.995	1e5	100
<i>Gamma</i>	<i>Tau</i>	<i>Learning Rate</i>	<i>Update time</i>	<i>Q_network (fc1)</i>	<i>Q_network (fc2)</i>	<i>Q_network (fc3)</i>
0.98	1e-3	5e-4	4	64	64	64

```

Episode 100    Average Score: 0.89
Episode 200    Average Score: 3.39
Episode 300    Average Score: 6.65
Episode 400    Average Score: 9.45
Episode 500    Average Score: 10.92
Episode 600    Average Score: 12.60
Episode 629    Average Score: 13.00
Environment solved in 529 episodes!    Average Score: 13.00

```



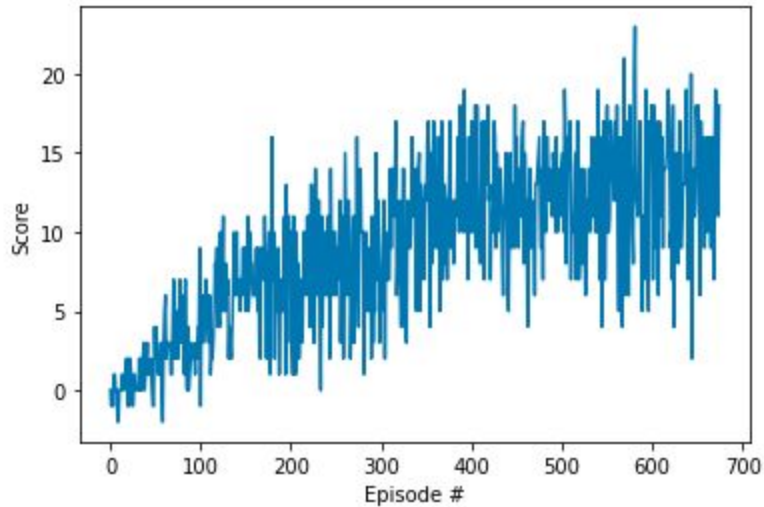
Experiment 2. Natural DQN. Twisting \max_t

$n_episode$	\max_t	eps_start	eps_end	eps_decay	$Buffer_Size$	$Batch_Size$
2000	3000	1.0	0.01	0.995	1e5	100
Γ	τ	Learning Rate	Update time	$Q_network (fc1)$	$Q_network (fc2)$	$Q_network (fc3)$
0.98	1e-3	5e-4	4	64	64	64

```

Episode 100    Average Score: 1.74
Episode 200    Average Score: 6.32
Episode 300    Average Score: 7.85
Episode 400    Average Score: 11.11
Episode 500    Average Score: 12.23
Episode 600    Average Score: 12.56
Episode 675    Average Score: 13.05
Environment solved in 575 episodes!    Average Score: 13.05

```



We check other Experiments to check behavior of the Q network dimension, and gamma parameter.

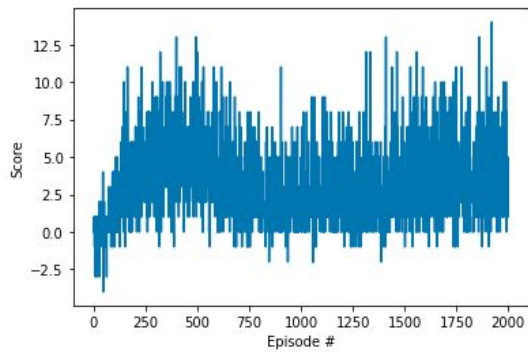
Experiment 3. Natural DQN. Twisting Neural Hidden Layer

<i>n_episode</i>	<i>max_t</i>	<i>eps_start</i>	<i>eps_end</i>	<i>eps_decay</i>	<i>Buffer_Size</i>	<i>Batch Size</i>
2000	1000	1.0	0.01	0.995	1e5	100
Gamma	Tau	Learning Rate	Update time	Q_network (fc1)	Q_network (fc2)	Q_network (fc3)
0.98	1e-3	5e-4	4	512	512	512

```

Episode 100    Average Score: 0.44
Episode 200    Average Score: 2.88
Episode 300    Average Score: 4.65
Episode 400    Average Score: 4.89
Episode 500    Average Score: 5.18
Episode 600    Average Score: 4.85
Episode 700    Average Score: 3.56
Episode 800    Average Score: 2.77
Episode 900    Average Score: 2.17
Episode 1000   Average Score: 2.20
Episode 1100   Average Score: 2.38
Episode 1200   Average Score: 2.70
Episode 1300   Average Score: 2.63
Episode 1400   Average Score: 2.95
Episode 1500   Average Score: 3.18
Episode 1600   Average Score: 3.54
Episode 1700   Average Score: 3.68
Episode 1800   Average Score: 3.64
Episode 1900   Average Score: 3.64
Episode 2000   Average Score: 3.82

```



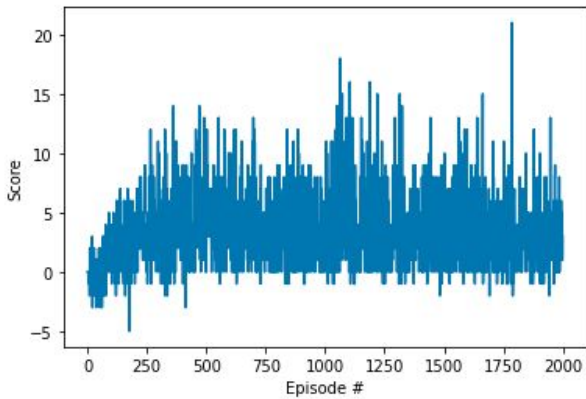
Experiment 4. Natural DQN. Twisting Gamma(Discount rate)

<i>n_episode</i>	<i>max_t</i>	<i>eps_start</i>	<i>eps_end</i>	<i>eps_decay</i>	<i>Buffer_Size</i>	<i>Batch Size</i>
2000	1000	1.0	0.01	0.995	1e5	100
Gamma	Tau	Learning Rate	Update time	Q_network (fc1)	Q_network (fc2)	Q_network (fc3)
0.8	1e-3	5e-4	4	64	64	64

```

Episode 100    Average Score: 0.45
Episode 200    Average Score: 2.13
Episode 300    Average Score: 3.48
Episode 400    Average Score: 4.08
Episode 500    Average Score: 4.39
Episode 600    Average Score: 4.64
Episode 700    Average Score: 3.89
Episode 800    Average Score: 3.37
Episode 900    Average Score: 3.56
Episode 1000   Average Score: 3.17
Episode 1100   Average Score: 4.57
Episode 1200   Average Score: 3.54
Episode 1300   Average Score: 3.53
Episode 1400   Average Score: 2.94
Episode 1500   Average Score: 3.76
Episode 1600   Average Score: 3.82
Episode 1700   Average Score: 2.98
Episode 1800   Average Score: 2.69
Episode 1900   Average Score: 2.70
Episode 2000   Average Score: 2.60

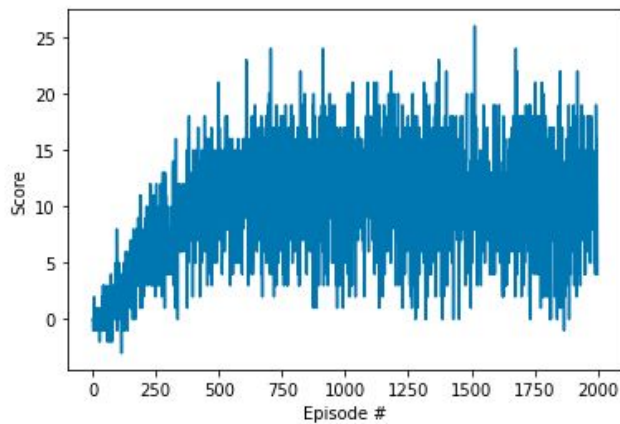
```



Experiment 5. Natural DQN. Increase Gamma(Discount rate)

<i>n_episode</i>	<i>max_t</i>	<i>eps_start</i>	<i>eps_end</i>	<i>eps_decay</i>	<i>Buffer_Size</i>	<i>Batch Size</i>
2000	1000	1.0	0.01	0.995	1e5	100
Gamma	Tau	Learning Rate	Update time	Q_network (fc1)	Q_network (fc2)	Q_network (fc3)
0.99	1e-3	5e-4	4	64	64	64

Episode 100	Average Score: 0.64
Episode 200	Average Score: 3.23
Episode 300	Average Score: 6.76
Episode 400	Average Score: 8.56
Episode 500	Average Score: 10.62
Episode 600	Average Score: 10.87
Episode 700	Average Score: 10.95
Episode 800	Average Score: 10.95
Episode 900	Average Score: 11.62
Episode 1000	Average Score: 11.76
Episode 1100	Average Score: 11.87
Episode 1200	Average Score: 11.93
Episode 1300	Average Score: 10.76
Episode 1400	Average Score: 11.10
Episode 1500	Average Score: 11.00
Episode 1600	Average Score: 11.11
Episode 1700	Average Score: 11.54
Episode 1800	Average Score: 11.58
Episode 1900	Average Score: 9.252
Episode 2000	Average Score: 11.28



Discussion: Some hyperparameters most effective to the DQN Algorithms are: Gamma (discount rate) if the discount rate $=0.99$ the agent does not get a good result and most depending all the future reward, so the score increases slowly. Otherwise, if the discount rate gamma $=0.8$ the agent does not get the high scores, may be declined for all of time learning. The best option of the gamma is 0.98 . The structure of the DQN network also affects the rewards of the agent, when we increase the number of nodes from 64 to 512, the result is worse. We can do more experiments to get the best result.