Q-learning

Q-learning is a reinforcement learning algorithm. It works by learning the function Q(s,a), where s and a stands for states and actions respectively. The final values of Q(s,a) determines the action-selection policy in every state. In order to train our Q(s,a), we have to decide 3 things, i.e., s (states), a (actions) and r(rewards). The function Q(s,a) is updated by the formula:

Q(state, action) = R(state, action) + Gamma \* Max[Q(next state, all actions)]

Where Gamma is the discount factor that controls the amount of influence that the future action have on the current state.

In the Q-learning, we have three important concepts: state, action and reward. When you are in one state, you can do one action to transfer to another state, and you will get the reward during this process.

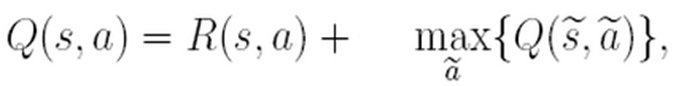
Our aim is to maximum the total reward. In our project, maximum the total reward means arriving the destination in the shortest time.

We have the reward matrix(R) at beginning. And the Q matrix a zero matrix.

(s represent state, a represent action)

In the R matrix and Q matrix, the column represent action and the row represent state.

The Q-Learning algorithm goes as follows:

1. Set the gamma parameter, and environment rewards in matrix R.
2. Initialize matrix Q to zero.
3. For each episode:
4. Select a random initial state.
5. Do while the goal state hasn’t been reached
6. Select one among all possible actions for the current state.
7. Using this possible action, consider going to the next state.
8. Get maximum Q value for this next state based on all possible actions.
9. 
10. Set the next state as the current state.

After our training, we can get the best order of Mario’s action by follow steps:

1. Set current state = initial state.
2. From current state, find the action with the highest Q value.
3. Set current state = next state.
4. Repeat Steps 2 and 3 until current state = goal state

The algorithm above will return the sequence of states from the initial state to the goal state.

REFERENCE: http://mnemstudio.org/path-finding-q-learning-tutorial.htm