**COMP7404 Assignment 2 Short Report**

SHI Zhongqi 3035347861

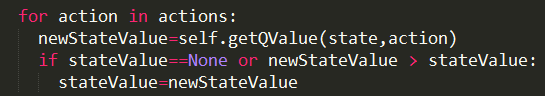
1. **Question 1**

In beginning state, all the values are set to None. The reward will give value for the exit state.

Then start iteration.

S is current state, a is the action to take, is next state. is the possibility. is discount. is the next state value.

  
 will check the Q(s,a) and take the maximum value.



1. **Question 2**

To make the agent wish to go right, the discount rate should be large. Just need to fulfill this equation : . So .

Together with the Noise rate: . So n should take as smaller as possible. Just give some number fulfill the above equation is ok.

Thus I give n=0.005 =0.95

1. **Question 3**

It is same with last question.

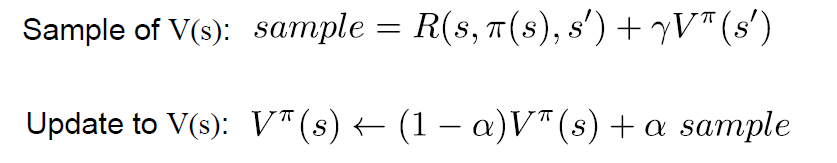
If agent wants the risking the cliff, then the noise should be small.

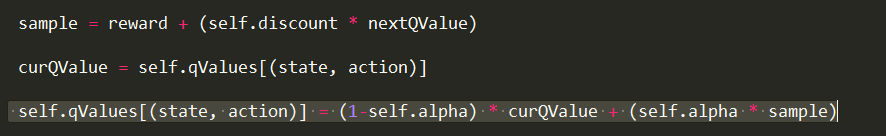
If agent wants higher exit value, then the discount rate should be large.

If want the agent live forever, then the living reward should be bigger than the exit value.

From the above statement, I give the value.

1. **Question 4**





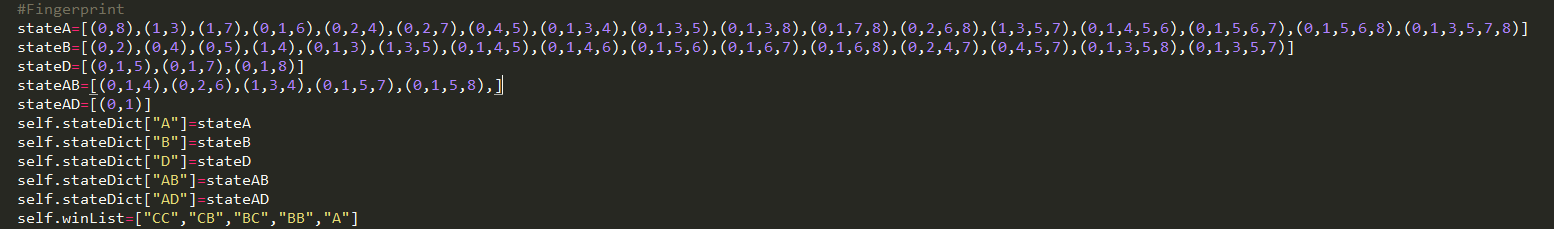
The QValue was saved as a diction using (state,action) as key. QValue return 0.0 if we have never seen a state or the Q node value.

Compute the action is just return the action that this state has the maximum QValue.

1. **Question 5**
2. **Question 6**

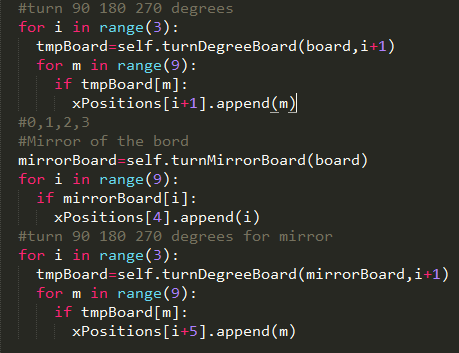
The key is the P-position and fingerprint.

Mark their coordinate and put them in a list is the easiest way to build the function of selection. Since C and CC is a bit special, so the judgement is done by the If and else condition.



In order to avoid the complicated status of different result for the timing, the string is used like “CC”, “CB”, “BC”, “BB”, “A”.

Then another difficult part is to do the mirror and turning for the fingerprint. Because it is really hard to mark all the finger print. So, I choose to do the computation.



First to get all the board status results, then check the actions, if the action could lead the board result to one of the P-Postion, then return it.