**Spring 2022 Introduction to Artificial Intelligence**

**Report of Homework #3**

Student name: 劉子齊 Jonathan

Student ID: 0716304

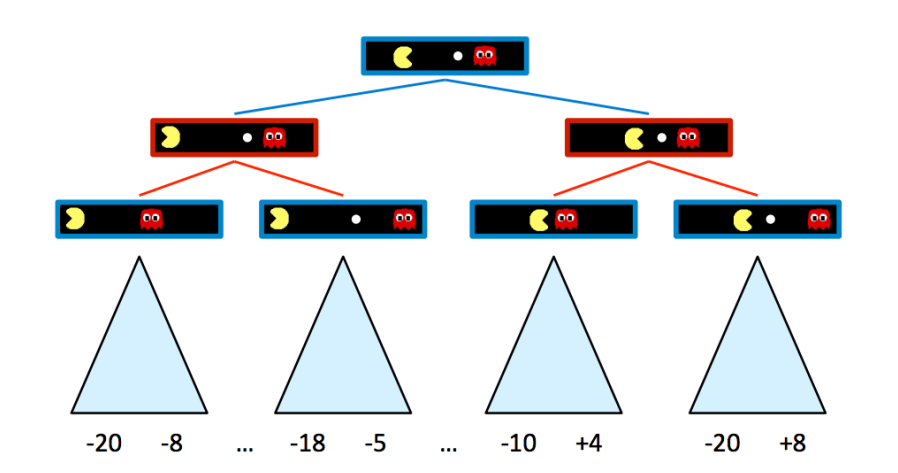
**Part 1 : Adversarial search**

Part 1-1: Minimax Search

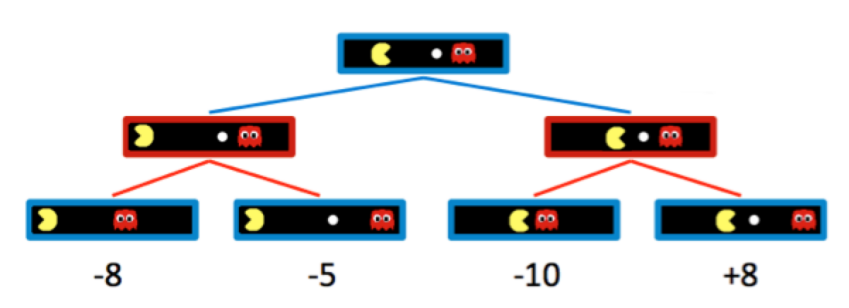
When the computer wants to make the next move in Pacman, of course it will play the one with the most points, which is eating all the food in this case, but this can easily fall into the opponent's trap, which is being eaten by the ghosts.

Therefore, a reasonable idea is to divide all levels into two categories of enemy and us, which will be Pacman and the Ghost in ours case. The more points the level under our side has, the better, and the level under the other side loses as few points as possible.

Also, we can't assume that the opponent is a fool, so at every level, we have to think that "the opponent may make the move that will cost us the most points", and we must choose the strategy of "minimizing maximum points loss" as much as possible. As a result, this strategy makes the Minimax Search.



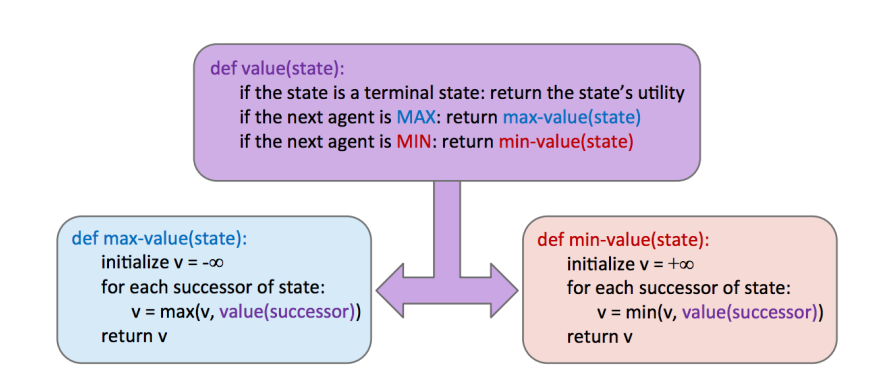
In the figure above, blue nodes correspond to nodes that Pacman controls and can decide what action to take, which will pick the choice with the max value, while red nodes correspond to ghost-controlled nodes, which will pick the minimum. Now let’s move onto the second depth of the tree in the figure above.

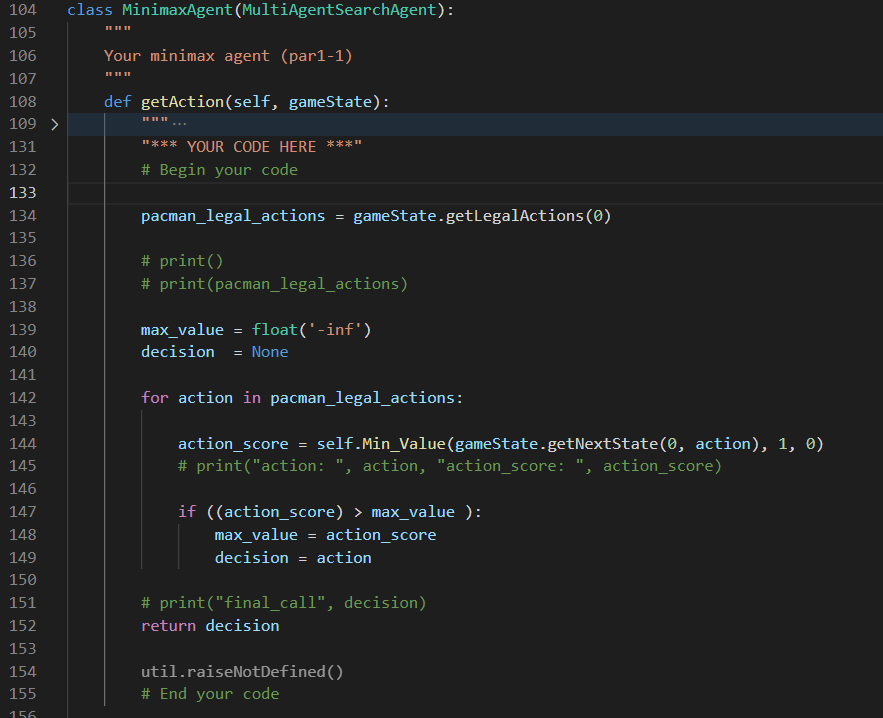


As we can observe from the figure above, the blue Pacman nodes chose the option with the max value, which Pacman believes to be the best choice. The minimax algorithm only maximizes over the children of nodes controlled by Pacman, while minimizing over the children of nodes controlled by ghosts. Hence, the two ghost nodes above have values of min(−8,−5) = −8 and min(−10,+8) = −10 respectively. Correspondingly, the root node controlled by Pacman has a value of max(−8,−10) = −8. As a result, the Pacman will get -8 as the score of this game.

However, if I was the Pacman in this game, I will mot be satisfied, since I knew that I could have get +8 as my final score, if I chose the right way. Hence, we will need to have the Pacman put some bet to move to way which is not so straightforward and will take some risks correspondingly.

When implementing the Minimax algorithm, I found it kind of similar to DFS (Deep First Search), which they both start with the leftmost terminal node and all the way to the right of the game tree. Basically, I followed the pseudocode found on a website of UC Berkeley in the figure below.





At the very first of my program in the figure above, I call the getLegalActions() function to get all the possible actions the Pacman can move for the next step. Besides, I declared the initial max\_value, which determines the final action, as a negative infinite float number, and the decision as a NULL.

After declaring, I construct a loop which depends on the elements in the list named “pacman\_legal\_actions”. For each action in the list, I first obtain the “action\_score” by the Minimax algorithm, then I compare all of them throughout the whole loop and get the final decision, which is the action with the highest “action\_score”.



For the Minimax algorithm, I separated it into two parts, one is for the minimum value, and the other for the maximum. For the Max\_Value() part, I first confirm if the current depth equals to the “self.depth”, which is the end of the game tree. If we’ve reach the end of the game tree, then I call the evaluationFunction() to get the final result and return it. Also, I’ll check whether there are no legal actions left, which might mean the Pacman just ran into a ghost and died. In this case, I’ll also return the final result obtained by the evaluationFunction().

If none of the two scenarios mentioned above happened, I’ll call the Min\_Value function, which will be mentioned later, and iterate it through all the legal actions obtained by the getLegalActions(0), which returns a list contains all the legal actions of Pacman since “0” stands for Pacman here. As I keep getting value from the Min\_Value function, I’ll compare them all and keep the maximum of all of them. Finally, I’ll return the maximum value.

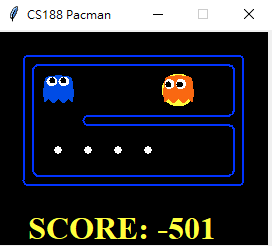


For the Min\_Value() part, I’ll check whether there are no legal actions left, which might means the Pacman just ran into a ghost and died. In this case, I’ll also return the final result of the current gameState obtained by the evaluationFunction().

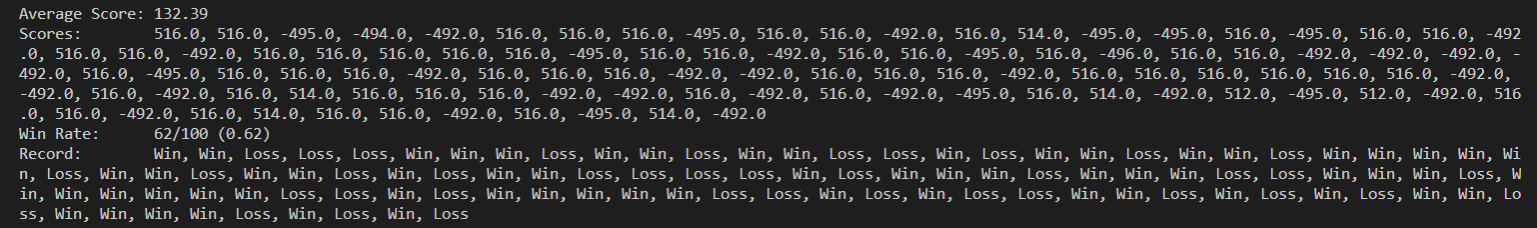
Then I’ll check if there’s still agents remaining. If there are still agents remaining, I’ll call the Min\_Value function, and iterate it through all the legal actions obtained by the getLegalActions(agentIndex), which returns a list contains all the legal actions of the corresponding agent. As I keep getting value from the Min\_Value function, I’ll sum them up and return the final value.

If there’s no agents remaining, I’ll call the Max\_Value function, and iterate it through all the depth and all the legal actions obtained by the getLegalActions(agentIndex), which returns a list contains all the legal actions of the corresponding agent. As I keep getting value from the Min\_Value function, I’ll compare them all and keep the minimum of all of them. Finally, I’ll return the minimum value.





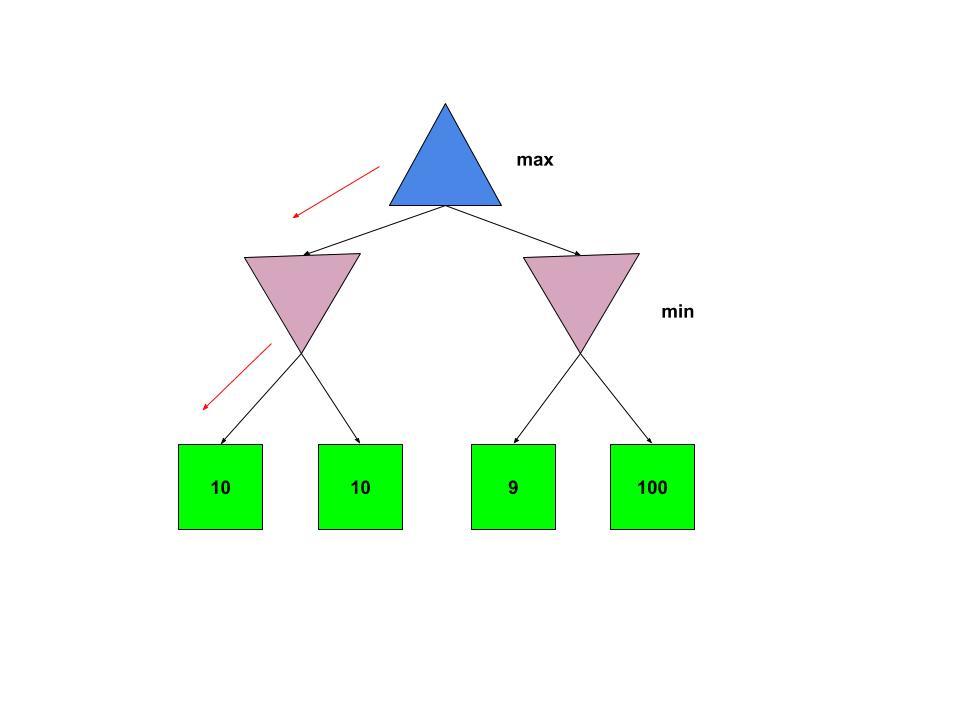
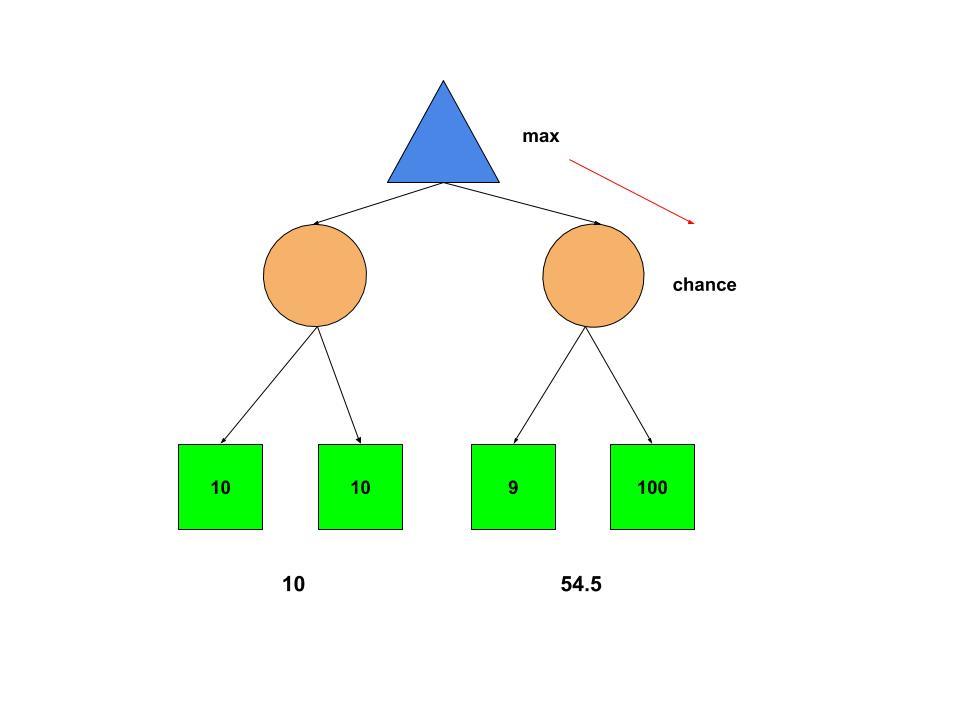




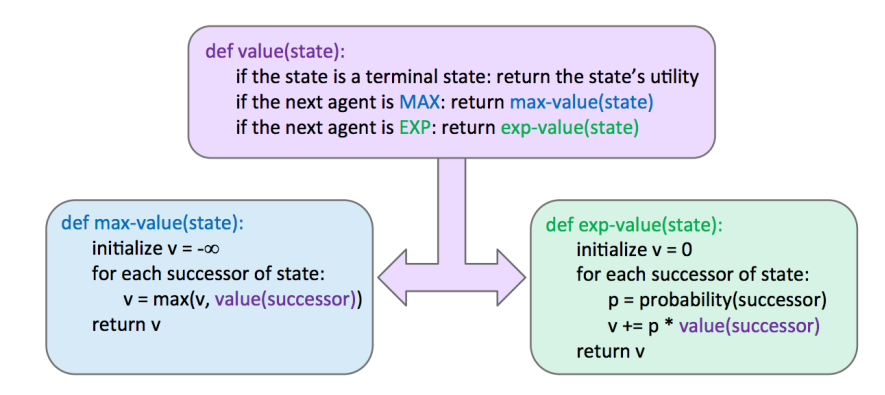
Part 1-2: Expectimax Search

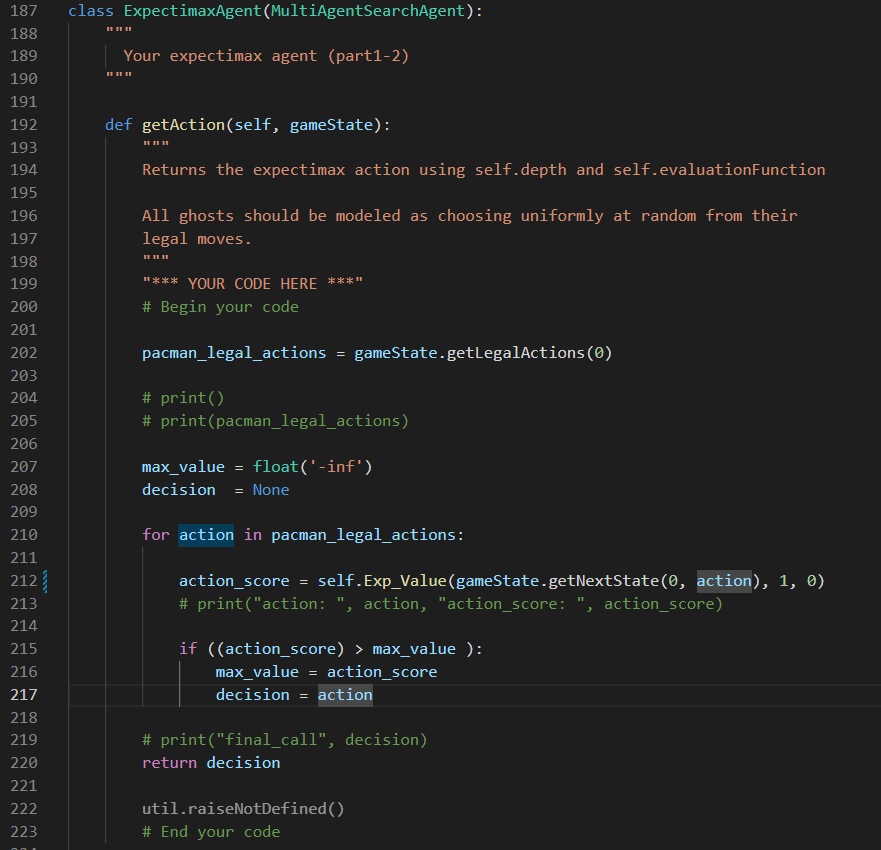
I’ll take Expectimax as a variation of Minimax Algorithm, which replaces minimizer nodes into chance nodes in the original Minimax game tree. As we know that the minimizer plays optimally, it makes sense to go to the left. But what if there is a possibility of the minimizer making a mistake. Therefore, going right might sound more appealing or may result in a better solution.

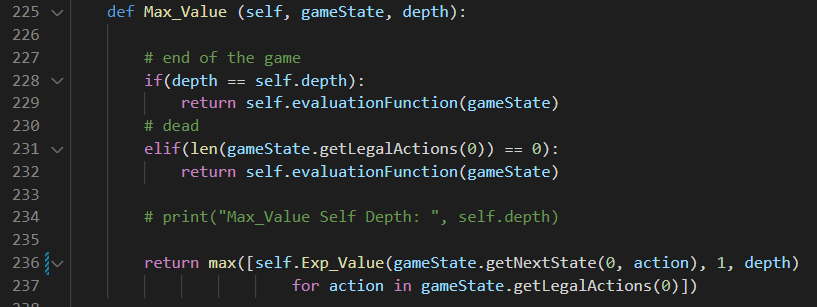
Hence, in the Expectimax Search, I replace minimizer nodes by chance nodes, which takes the average of the agent numbers, which can be seen in the figure below. On the left, we have the original Minimax Search, and on the left we have the tree of Expectimax Search.

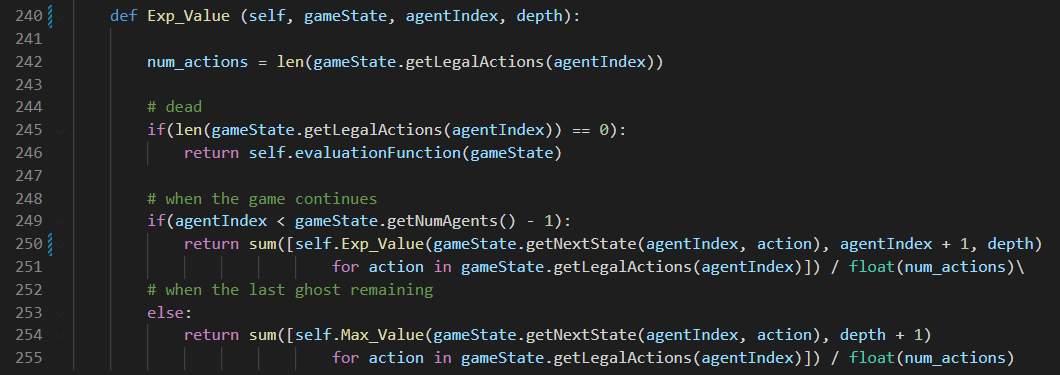
For the implementation, I followed the pseudocode found on a website of UC Berkeley in the figure below. Which we can find both of them are basically the same, except for changing the Min\_Value() function of the Minimax Search into Exp\_Value() function which contains the chance node, of the Expectimax Search.







By the figure, we can find the get\_action() function and the Max\_Value() function are almost the same. I also checked the same condition and return the same value in both functions as I did in the Minimax Search. The only thing changed was in the Max\_Value() function, which I changed the original Min\_Value() function into Exp\_Value() function.

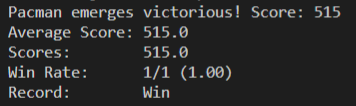


In the Exp\_Value() function, basically, it’s also kind of similar to the Min\_Vlaue() function of the Minimax Search. I checked whether there are no legal actions left, which might mean the Pacman just ran into a ghost and died. In this case, I’ll also return the final result of the current gameState obtained by the evaluationFunction().

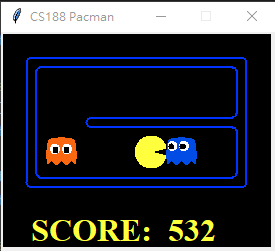
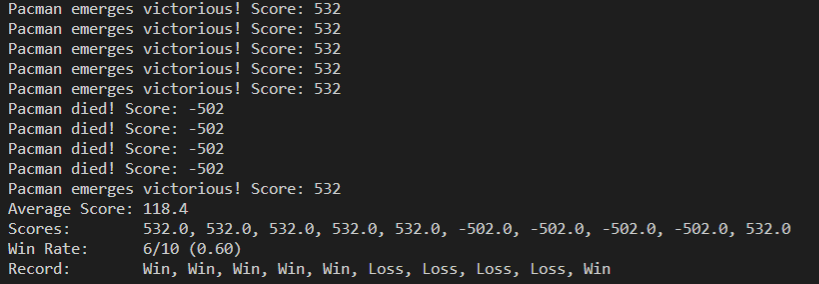
Then I’ll check if there’s still agents remaining. If there are still agents remaining, I’ll call the Exp\_Value() function, and iterate it through all the legal actions obtained by the getLegalActions(agentIndex), which returns a list contains all the legal actions of the corresponding agent. As I keep getting value from the Exp\_Value() function, I’ll sum them up. After getting the final sum value, I divided the sum with the total num of the actions and return the final value.

If there’s no agents remaining, I’ll call the Max\_Value() function, and iterate it through all the depth and all the legal actions obtained by the getLegalActions(agentIndex), which returns a list contains all the legal actions of the corresponding agent. As I keep getting value from the Max\_Value() function, I’ll sum them up. After getting the final sum value, I divided the sum with the total num of the actions. Finally, I’ll return the final value.





Part 1-3: Evaluation Function (Bonus)

[n3.pdf (berkeley.edu)](https://inst.eecs.berkeley.edu/~cs188/fa18/assets/notes/n3.pdf)

**Part 2 : Q-learning**

Part 2-1: Value Iteration

Part 2-2: Q-learning

Part 2-3: epsilon-greedy action selection

Part 2-4: Approximate Q-learning

**Part 3 : DQN**

**Part 4 : Try other SOTA methods (Bonus)**