Homework 2 - Choice of Indiana Jones

Jan 12, 2021

In this homework assignment you are going to calculate the expected utility of an action in a partially observable environment.

Background Story

You are Dr. Indiana Jones, an archaeologist and adventurer. When you are searching a lost ancient temple, you fell into a circular corridor through a trap door on the ceiling.

Fortunately, you have a map of this corridor (shown in Figure 1). There are six doors leading to six different rooms, clockwisely marked on the map from 0 to 5. In room 0, you can find the holy grail, which has a utility of 2000. In room 1 there are poisonous flowers which have a utility of -500. In room 2 are silver coins with a utility of 100. In room 3 you can find a golden idol which has a utility of 1000. Room 4 is an empty room which has a utility of 0. In room 5 there are snakes which have a utility of -5000.

However, the 6 doors look exactly the same. You are now in front of a door, but you do not know which door it is. By observing your environment, you can get the probability of which door you are in front of. That is, $P(S_0|e)$={Door 0: 0.125, Door 1: 0.25, Door 2: 0.0625, Door 3: 0.0625, Door 4: 0.25, Door 5: 0.25}.

You have to choose one door to pass through, then exit the temple from the other side of the room. Your choices are {0-open the door in front of you; 1-open the door which is the first one clockwise; 2-open the door which is the second one clockwise; 3-open the door which is the third one clockwise; 4-open the door which is the fourth one clockwise; 5-open the door which is the fifth one clockwise}. Remember you can only choose one door to open.

However, you have to enter a rotating platform in the center before you jump to your target door if your target door is not the one in front of you. When you jump, you have a 90% chance to land in front of the target door. You have a 5% chance to land at the door left to your target door. You have a 5% chance to land at the door right to your target door. If you choose to open the door in front of, you have a 100% chance to open it. Once you jump to a door from the platform or open a door, you have to enter the room and cannot open other doors anymore.

Your task is to calculate the expected utility of each action so that you can choose the action that maximizes the expected utility.

Preparation - Expected value of a function of a random variable

The formula is

$$E(f(X)) = \sum_{i=1}^{n} p(x_i) f(x_i)$$

You are supposed to complete the function **expect** in the *utility.py* file.

There are two input parameters of the function **expect**. (1) A dictionary representing the distribution of the random variable. The keys of dictionary are the possible values of random variable X and the values of the dictionary are the corresponding probability $p(x)$. (2) A function. The function f accepts a scalar as input and returns a scalar.

The return value of the function **expect** is a scalar, representing the expected value of $f(X)$.

You can use the expect function from Homework 1.

Task 1 - Probability of the next state after you take action

The formula is

$$P(result(a) = s'|a, e) = \sum_{s} P(result(s, a) = s'|a) P(S_0 = s|e)$$

.

You are supposed to complete the **getSPrimeProbability** function in the *utility.py* file. You should use the **expect** function.

There are four input parameters of the function **getSPrimeProbability**. (1) sPrime - a scalar representing the next state as a result of your action. (2) action - a scalar representing which action you are taking. (3) s0Distribution - a dictionary representing the probability distribution of the present state (S_0) given evidence. The keys of this dictionary are all possible current states. The values of this dictionary are their corresponding probability ($P(S_0 = s|e)$). (4) transitionTable - a dictionary showing the possible results of actions. The keys of this dictionary are all possible current states. The values of this dictionary are action dictionaries whose keys represent all possible actions. The values of the action dictionaries are probability dictionaries whose keys are all possible results from the action given the current state. The values of the probability dictionaries are the probability of the results given the action and the current state.($\{s_0 : \{a : \{s' : P(Result(s,a) = s'|a)\}\}\}$)

The return value of the function **getSPrimeProbability** is a scalar. It represents the probability of you ending up in the state of sPrime given your action.

Task 2 - Expected utility of an action

The formula is

$$EU(a|e) = \sum_{s'} P(result(a) = s'|a,e)U(s')$$

You are supposed to complete the **getEU** function in the *utility.py* file. You should use the **expect** function.

There are two input parameters of the function **getEU**. (1) sPrimeDistributionGivenAction - a dictionary representing the probability distribution of the next state given your action ($P(Result(a) = S'|a,e)$). The keys of this dictionary are all possible next states. The values of this dictionary are their corresponding probability given your action. (2) utilityTable - a dictionary representing the utility of each state. The keys of this dictionary are all possible states. The values of this dictionary are their corresponding utility.

The return value of the function **getEU** is a scalar, representing the expected utility of a certain action.

Task 3 - Main function

You are supposed to complete the **main** function in the *utility.py* file.

In the main function, you are provided with all the information you need: the distribution of S_0 given evidence, the utility table of each state, and the transition table $\{s_0 : \{a : \{s' : P(Result(s,a) = s'|a)\}\}\}$.

You need to complete the two dictionaries. The keys of the first dictionary (**sPrimeDistribution**) are all possible actions. The values are dictionaries whose keys are all possible S' given the action and values are the corresponding probability $P(s'|a,e)$. ($\{a : \{s' : P(s'|a,e)\}\}$). You will need the function **getSPrimeProbability** to complete the **sPrimeDistribution** dictionary. The keys of the second dictionary (**EU**) are all possible actions. The values are the expected utility of the action. ($\{a : EU(a|e)\}$). You will need the function **getEU** to complete the **EU** dictionary.

Test examples

This example has three parameters: (1) s0Distribution - a dictionary representing the probability distribution of the present state (S_0) given evidence. The keys of this dictionary are all possible current states. The values of this dictionary are their corresponding probability ($P(S_0 = s|e)$). (2) utilityTable - a dictionary representing the utility of each state. The keys of this dictionary are all possible states. The values of this dictionary are their corresponding utility. (3) transitionTable - a dictionary showing the possible results of actions. The keys of this dictionary are all possible current states. The values of this dictionary are action dictionaries whose keys represent all possible actions. The values of the action dictionaries are probability dictionaries whose keys are all possible results from the action given the current state. The values of the probability dictionaries are the probability of the results given the action and the current state. $(\{s_0 : \{a : \{s' : P(Result(s, a) = s'|a)\}\}\})$

For this problem, your code should print out the expected utility dictionary (approximately) {0: -1056.25, 1: -689.0624999999999, 2: 248.12499999999994, 3: -174.06250000000003, 4: -796.25, 5: -6.875}.

Submission

Please submit a completed *utility_YourLastName_YourFirstName.py* file on CCLE before due. **The due date and time of this homework assignment is Monday, 01/25/2021 11:59pm**.

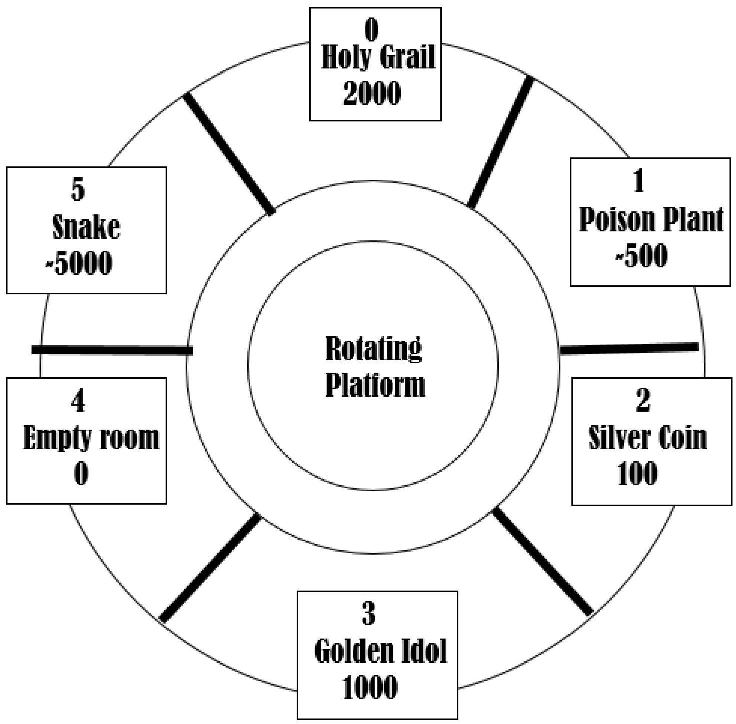

Figure 1: Map of the corridor