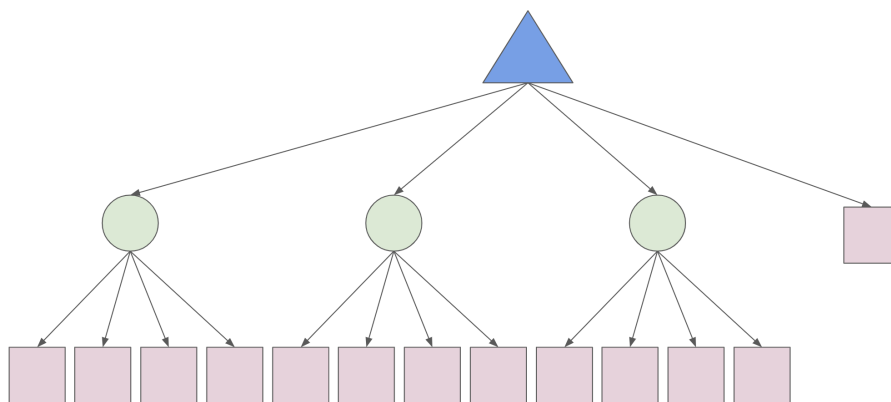


Q1. [30 pts] Expectimax Yahtzee

Consider a simplified version of the game *Yahtzee*. In this game, we have 3 regular tetrahedral dice with 4 sides each (numbered 1-4) and the game begins by rolling all 3 dice. At this point, a player can make a decision: pick one of the 3 dice to reroll, or don't reroll anything. Then, points are assigned as follows:

- A reward of 10 points is given for two-of-a-kind (for example, 4-4).
- A reward of 15 is given to three-of-a-kind (for example, 4-4-4).
- A reward of 7 points is given for rolling a series (1-2-3 or 2-3-4).
- Otherwise (or if the sum is higher than the special reward), the score is equal to the sum of all 3 dice.

(a) We will formulate this problem as an expectimax tree. The resulting tree for the problem is drawn below.



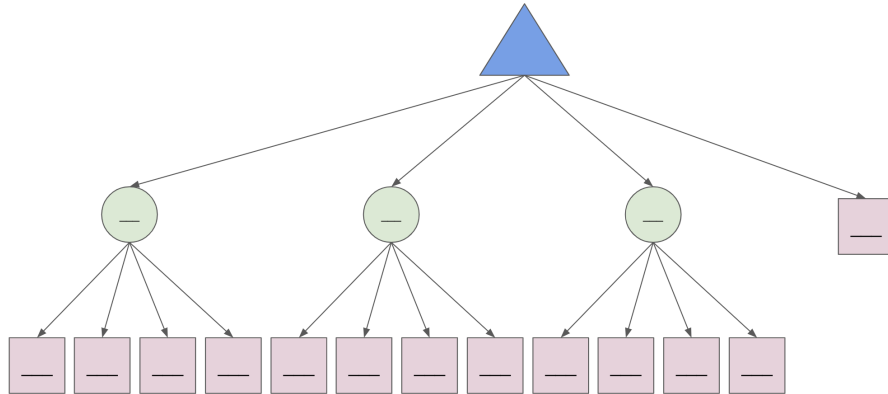
(i) [1 pt] The resulting tree for the problem is drawn below. Which part of the game does the maximizer node represent?

The player can make a decision: pick one of the 3 dice to reroll, or don't reroll anything. (The maximizer node represents the player's choice of re-rolling dice.)

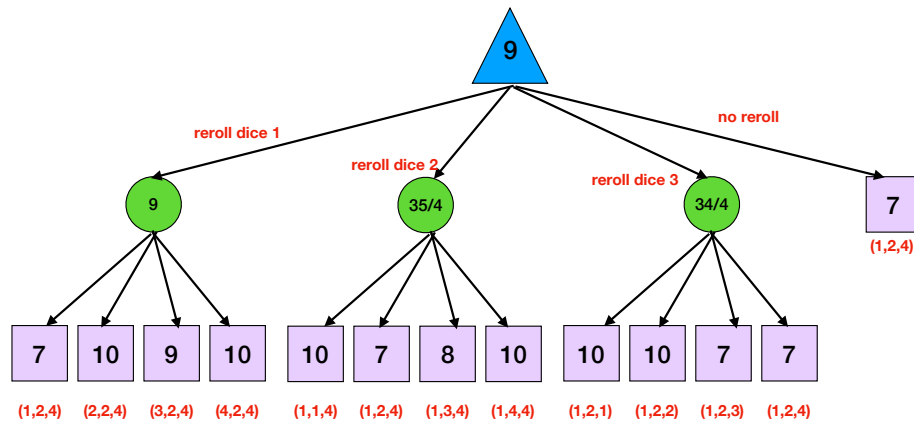
(ii) [1 pt] Which part of the game do the chance nodes represent?

The selected dice is rerolled. (The chance nodes represent the outcome of the dice rolls.)

- (b) Suppose that the initial roll of the 3 dice results in (1, 2, 4), and we want to know what action the player should take. For this question, we recommend filling in the expectimax tree on scratch paper before entering your answers.



- (i) [1 pt] Which value appears in the terminal node at depth 2 (the direct child of the root)?
 7, The right-most branch corresponds to the "no re-roll" action, since there's no dice roll chance node and the game ends immediately. If we don't re-roll, we end up with a score of 7.
- (ii) [3 pts] Which three values appear in the chance nodes of the expectimax tree?
 8.5, 8.75, 9

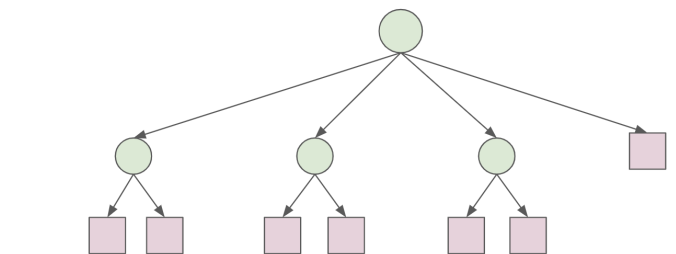
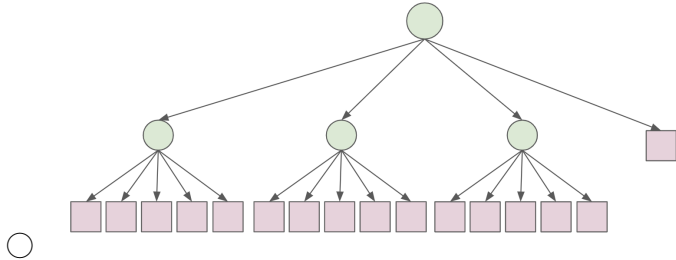
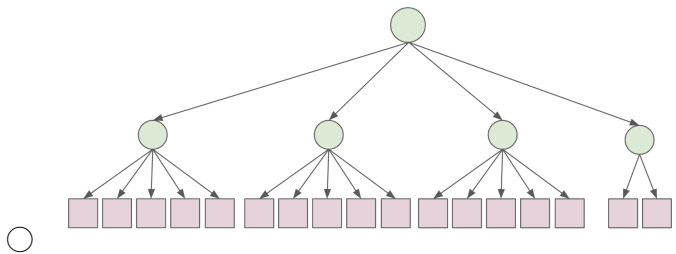
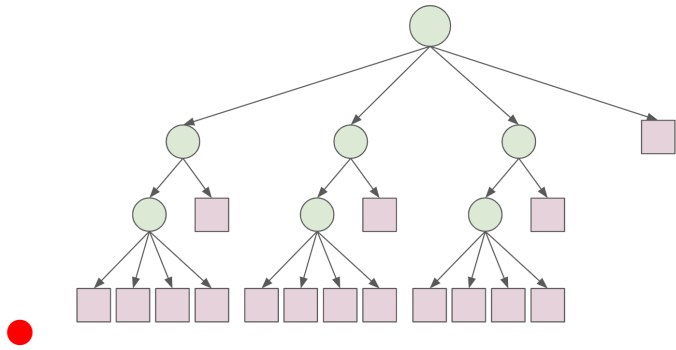


- (iii) [1 pt] Which action should the player take?
 Reroll die 1, Based on the solution expectimax tree, the best action for the agent to take is "reroll die 1" since 9 is a greater expected value than that of the other actions.

Now suppose the human player does not understand how to play the game, and as a result, they choose any action with uniform probability, regardless of the initial roll. Moreover, we assume that the human's choice will be carried out by a "somewhat helpful" robot called Albertbot: given a configuration of dice and the desired action from the human, this robot either actually implements the human's action (with probability $1 - p$) or overrides it with a 'no reroll' action (with probability $p > 0$). If the human action is already 'no reroll', then the robot does not interfere.

(c) Given a particular Yahtzee roll *roll*, Let A , B , C and D be the expected reward of performing actions 'reroll die 1', 'reroll die 2', 'reroll die 3', and 'no reroll', respectively.

(i) [1 pt] Which of the following trees best represents the expectimax tree, after accounting for the presence of the Albertbot?



None of the above.

(ii) [5 pts] Express R_H and R_{AH} in terms of A , B , C and D , where:

- R_H is the expected reward for the human acting *without* Albertbot's help.
- R_{AH} is the expected reward for the human acting *with* Albertbot's help.

Show all steps of your work and write your expression into the form of $X + Yp$, where X and Y are expressions that contain A , B , C and D but not p .

$$R_H = \frac{A+B+C+D}{4}$$

$$R_{AH} = \frac{(A+B+C)(1-p)}{4} + \left(\frac{3p}{4} + \frac{1}{4}\right)D = \frac{A+B+C}{4} - \frac{(A+B+C)p}{4} + \frac{3D}{4}p + \frac{D}{4} = \frac{A+B+C}{4} + \frac{D}{4} - \frac{A+B+C}{4}p + \frac{3D}{4}p$$

$$= \frac{A+B+C+D}{4} + \frac{3D-(A+B+C)}{4}p$$

(iii) [3 pts] What is the condition for our Albertbot to strictly increase expected reward? Write the condition above using only $A, B, C, D, >$.

In part (ii) we found that $R_{AH} = R_H + Yp$ where $Y = \frac{3D-(A+B+C)}{4}$. Therefore, for Albertbot to strictly increase the expected reward, we need $Y > 0$, shown as:

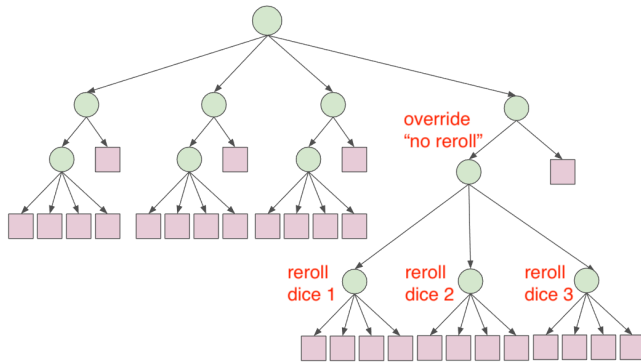
$$3D - (A + B + C) > 0 \text{ or } D > \frac{A + B + C}{3}$$

(iv) [2 pts] In one sentence, please describe the situation when the condition above is true.

The above statement shows that the no-reroll choice (D) gives a higher reward than the average of rerolling one of the dice.

(d) Your friend Diana argues that a helpful robot should not only override the human player's "reroll" choice with probability p (and replace it with a "no reroll"), but also override the human player's "no reroll" choice with probability p (and replace it with the outcome of selecting one of the 3 dice at random and rerolling that dice). Diana needs your help with drawing the new expectimax tree for the Dianabot.

(i) [5 pts] Draw the expectimax tree for Dianabot. You need to draw out the tree with all nodes in their correct shapes; you do not need to label any values in the tree. Hint: you can start by modifying the expectimax tree for the Albertbot.



- (ii) [5 pts] What is the expected reward for a random human player with Dianabot's assistance? Again, please **show all steps of your work** and **write your expression into the form of $X + Yp$** , where X and Y are expressions that contain A , B , C and D but not p .

$$\begin{aligned}R_{DH} &= \frac{1}{4}[(1-p)(A+B+C)] + \frac{3}{4}pD + \frac{1}{4}(1-p)D + \frac{1}{4} \times \frac{1}{3}p(A+B+C) \\ &= \frac{3-2p}{12}(A+B+C+D) + \frac{2}{3}pD \\ &= \frac{1}{4}(A+B+C+D) + \frac{1}{6}(3D-A-B-C)p\end{aligned}$$

- (iii) [2 pts] Under what condition on A , B , C , D is the human better off using Dianabot rather than Albertbot? Comparing the two expressions, we see that both are helpful when $D > (A+B+C)/3$, but Albertbot is more helpful than Dianabot (because $1/4 > 1/6$); both are unhelpful when $D < (A+B+C)/3$, but Albertbot is more unhelpful than Dianabot (for the same reason). Thus, the human is better off choosing Dianabot when $D < (A+B+C)/3$; but of course, in that case, the human is better off choosing no robot at all.