

**Due:** Wednesday, February 9 at 10:59pm (submit via Gradescope).

**Policy:** Can be solved in groups (acknowledge collaborators) but must be written up individually

**Submission:** Your submission should be a PDF that matches this template. Each page of the PDF should align with the corresponding page of the template (page 1 has name/collaborators, question 1 begins on page 2, etc.). **Do not reorder, split, combine, or add extra pages.** The intention is that you print out the template, write on the page in pen/pencil, and then scan or take pictures of the pages to make your submission. You may also fill out this template digitally (e.g. using a tablet.)

First name	<i>Qingjing</i>
Last name	<i>Zhang</i>
SID	<i>3037581676</i>
Collaborators	<i>None</i>

**For staff use only:**

Q1.	Expectimax Yahtzee	/40
	Total	/40

## Q1. [40 pts] Expectimax Yahtzee

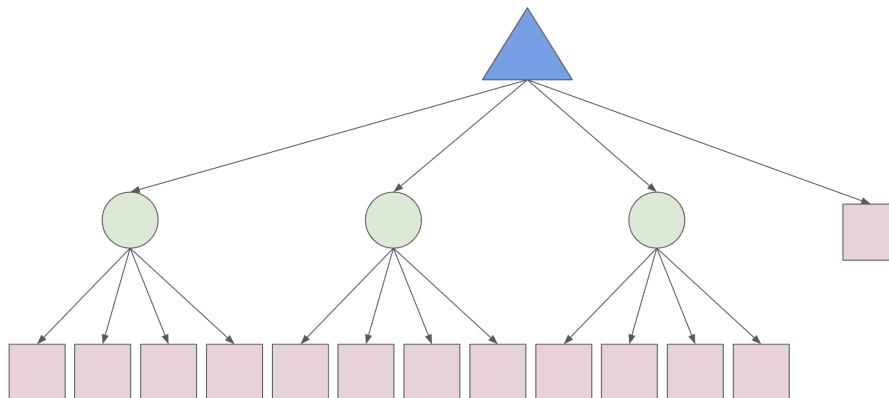
Consider a simplified version of the game *Yahtzee*. In this game, we have 3 regular tetrahedral dice with 4 sides each (numbered 1-4) and the game begins by rolling all 3 dice. At this point, a player can make a decision: pick one of the 3 dice to reroll, or don't reroll anything. Then, points are assigned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

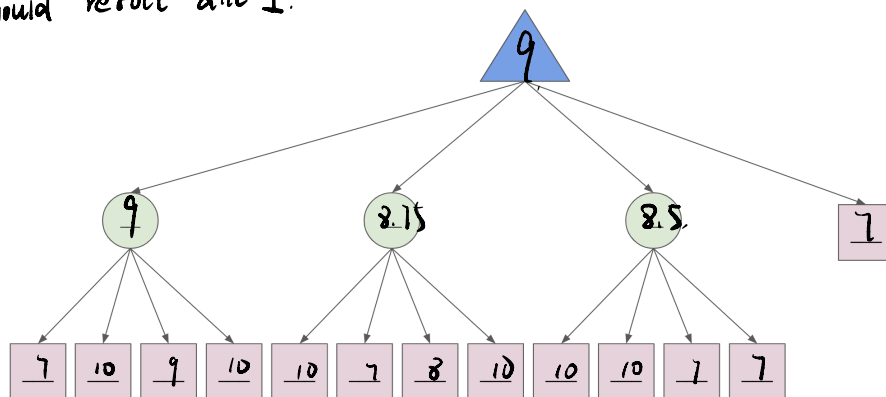
(i) [3 pts] The resulting tree for the problem is drawn below. Given a specific initial roll, the branching factor (of the player's decision) from the root node is 4. The branching factor at the chance nodes is 4. What do those chance nodes represent? (There are multiple solutions, you only need to write down one solution)

- Chance node 1: Expected reward after rerolling die 1
- Chance node 2: Expected reward after rerolling die 2
- Chance node 3: Expected reward after rerolling die 3.



- (ii) [7 pts] Given a starting roll (1,2,4) (corresponding to the outcomes of die rolls 1, 2, and 3 respectively), what move should you take? Fill in the values of the expectimax tree below to justify your answer.

we should reroll die 1.



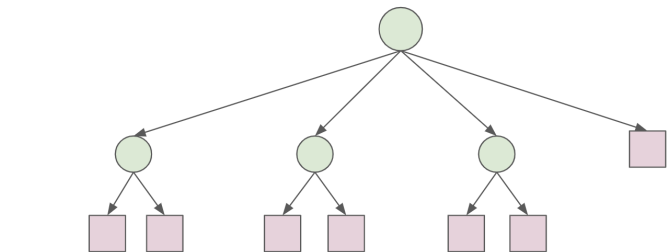
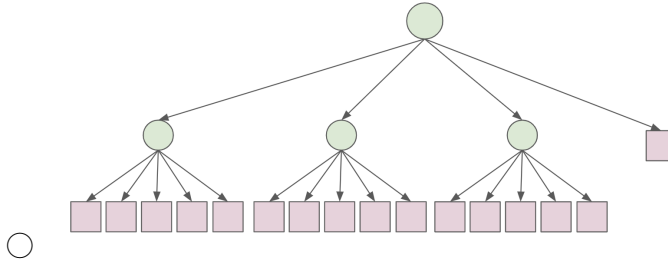
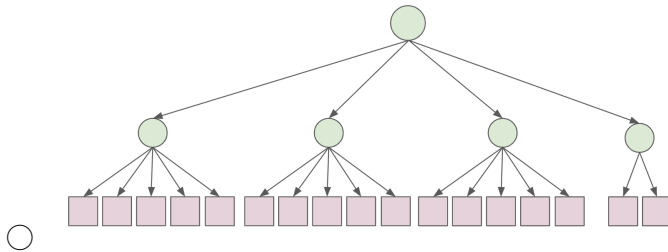
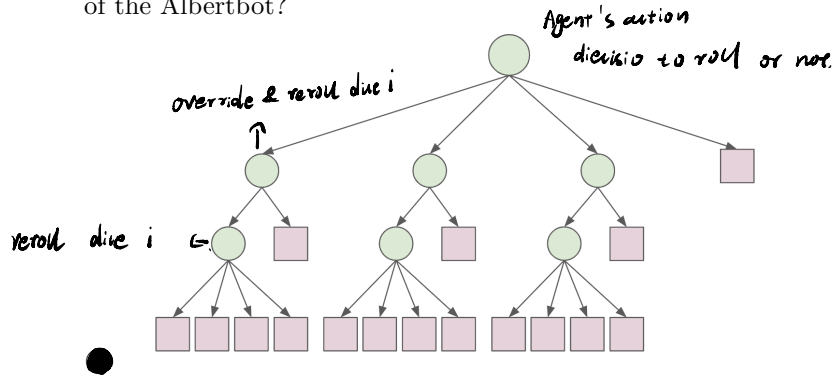
draft:

1	2	4	7	1	1	4	10	1	2	1	10
2	2	4	10	1	2	4	7	1	2	2	10
3	2	4	9	1	3	4	8	1	2	3	7
4	2	4	10	1	4	4	10	1	2	4	7

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.

(b) 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) [3 pts] 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  $p$ .

$$R_H = \frac{1}{4} (A+B+C+D) \\ = \frac{1}{4} (A+B+C+D)$$

$$R_{AH} = \frac{1}{4} [(1-p)(A+B+C+D) + 4pD] \\ = \frac{1}{4} (A+B+C+D) + \frac{1}{4} (3D-A-B-C)p.$$

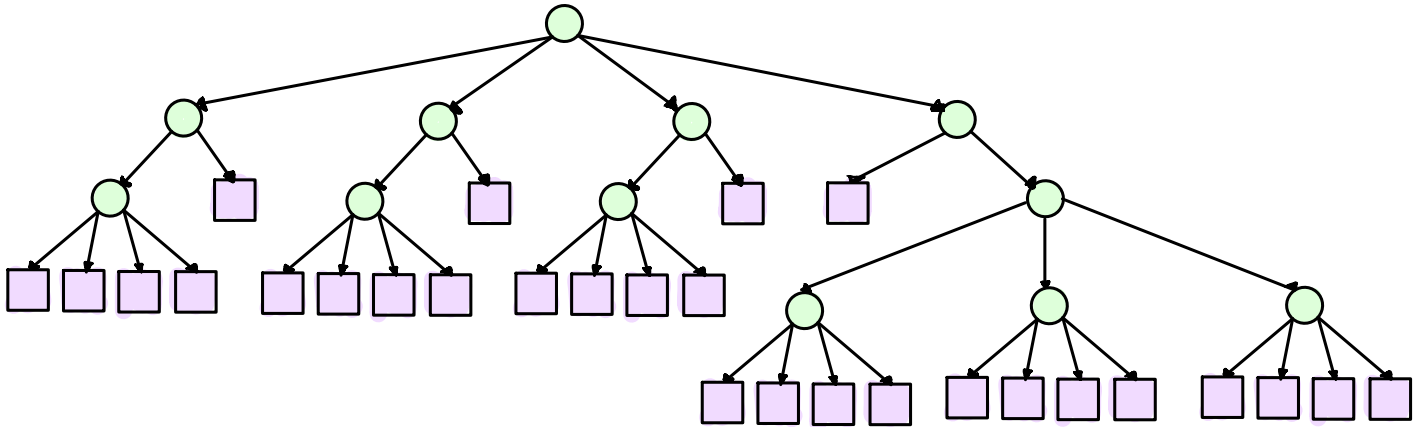
(iii) [5 pts] What is the condition for our Albertbot to strictly increase expected reward? Write the condition above using only  $A$ ,  $B$ ,  $C$ ,  $D$ ,  $>$ .

$$R_{AH} > R_H \\ \frac{1}{4} (A+B+C+D) + \frac{1}{4} (3D-A-B-C)p > \frac{1}{4} (A+B+C+D) \\ \Rightarrow 3D > A+B+C. \\ 3 > \frac{A+B+C}{D}$$

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

The average reward of re-rolling die 1, 2, 3  
is greater than the average reward of no re-rolling.

- (c) 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  $p$ .

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

- (iii) [5 pts] Under what condition on  $A, B, C, D$  is the human better off using Dianabot rather than Albertbot?

$$R_{DH} > R_{AH}$$

$$\frac{1}{4} (A+B+C+D) + \frac{1}{4} [2pD - \frac{2}{3}(A+B+C)p] > \frac{1}{4} (A+B+C+D) + \frac{1}{4} (3D-A-B-C)p$$

$$\begin{aligned}
 \Rightarrow \frac{1}{4} [2pD - \frac{2}{3}(A+B+C)p] &> \frac{1}{4} (3pD - A - B - C)p \\
 (A+B+C) &> 3D
 \end{aligned}$$

Under the condition that  $A+B+C > 3D$ ,  
Dianabot is better than Albertbot.