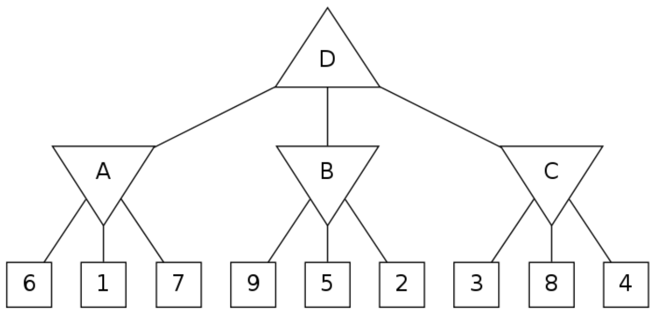
1. (7 points) Minimax

Consider the zero-sum game tree shown below. Triangles that point up, such as at the top node (root), represent choices for the maximizing player; triangles that point down represent choices for the minimizing player. Outcome values for the maximizing player are listed for each leaf node, represented by the values in squares at the bottom of the tree. Assuming both players act optimally, carry out the minimax search algorithm. Enter the values for the letter nodes in the boxes below the tree.




A: ***1*** B: ***2*** C: ***3*** D: ***3***

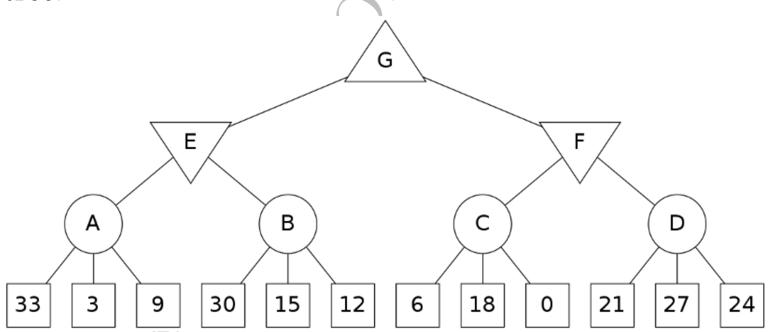
Explanation. Maximizing nodes choose the highest value from their children, and minimizing node take on the lowest value from among their children.

A chooses 1, B chooses 2, and C chooses 3.

D then chooses 3 from among A, B, and C.

2. (8 points) Expectiminimax

Consider the game tree shown below. As in the previous problem, triangles that point up, such as the top node (root), represent choices for the maximizing player; triangles that point down represent choices for the minimizing player. The circular nodes represent chance nodes in which each of the possible actions may be taken with equal probability. The square nodes at the bottom represent leaf nodes. Assuming both players act optimally, carry out the expectiminimax search algorithm. Enter the values for the letter nodes in the boxes below the tree.




A: ***15*** B: ***19*** C: ***8*** D: ***24***

E: ***15*** F: ***8*** G: ***15***

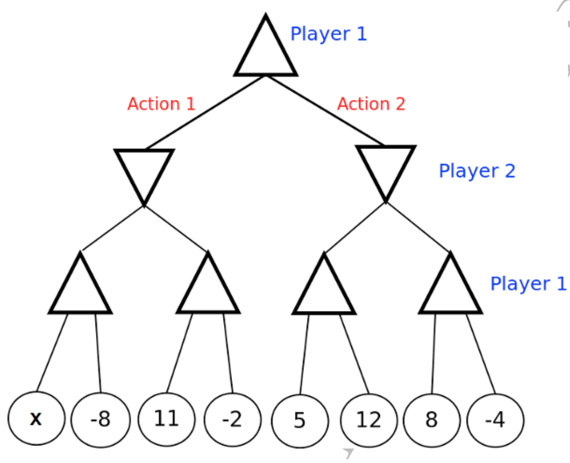
Explanation. The value for each circular node is equal to the expectation of the values of its children, which is found by adding up each of the child values and dividing by the number of children.

E and F take the minimums from their respective children.

G takes 15, the max of E and F

3. (12 points) UNKNOWN LEAF VALUE

Consider the game tree in the figure below, where one of the leaves has an unknown payoff, x. Player 1 moves first, and attempts to maximize the value of the game.






Each of the next 3 questions asks you to write a constraint on x specifying the set of values it can take. Please specify your answer in one of the following forms:

- Write "All" if x can take on all values.
- Write "None" if x has no possible values.
- Use an inequality in the form "x<value", "x>value", or "value1<x< value2" to specify an interval of values. As an example, if you think x can take on all values larger than 16, you should enter "x>16".

(a) Assume Player 2 is a minimizing agent (and Player 1 knows this). For what values of x is Player 1 guaranteed to choose Action 1 for their first move?

*Answer.* $x > 8$

Explanation. Depending on the value of x, there can be 3 different values for taking Action 1:

If $x < -8$, then Action 1 results in -8;

If $-8 \leq x \leq 11$ then Action 1 results in $x$;

If $x > 11$ then Action 1 results in 11.

Action 2 always results in a utility of 8. Hence Action 1 is optimal for Player 1 if $x > 8$.

(b) Assume Player 2 chooses actions at random with each action having equal probability (and Player 1 knows this). For what values of x is Player 1 guaranteed to choose Action 1?

*Answer.* $x > 9$

Explanation. Action 2 gives Player 1 a utility of 10, so the average of $x$ and 11 must be greater than 10. $(x + 11)/2 > 10 \rightarrow x > 9$

(c) Denote the minimax value of the tree as the value of the root when Player 1 is the maximizer and Player 2 is the minimizer. Denote the expectimax value of the tree as the value of the root when Player 1 is the maximizer and Player 2 chooses actions at random (with equal probability). For what values of x is the minimax value of the tree worth more than the expectimax value of the tree?

*Answer:* ***(None)***

Explanation. To satisfy we would need $x$ such that $\max(\min(x, 11), 8) > \max((x + 11)/2, 10)$.
For $x \leq 8$: minimax has value 8; expectimax has value 10 since $(x + 11)/2 < 10$; $8 \not> 10$.
For $8 < x \leq 9$: minimax has value $x$; expectimax has value 10 since $(x + 11)/2 \leq 10$; $x \not> 10$.
For $9 < x \leq 11$: minimax has value $x$; expectimax has value $(x + 11)/2$; $x \not> (x + 11)/2$.
For $11 < x$: minimax has value 11; expectimax has value $(x + 11)/2$; $11 \not> (x + 11)/2$.

(d) Is it possible to have a game, where the minimax value is strictly larger than the expectimax value? ☐ Yes √ ***No***

Explanation. The minimax value can never be strictly greater than the expectimax value for the same tree because in minimax Player 2 always chooses the worst possible move for Player 1, while in expectimax, those same nodes average that value with other higher values. Thus, the utility at a node under expectimax is always at least as high as the utility of the same node under minimax.

4. (9 points) Alpha-Beta Pruning

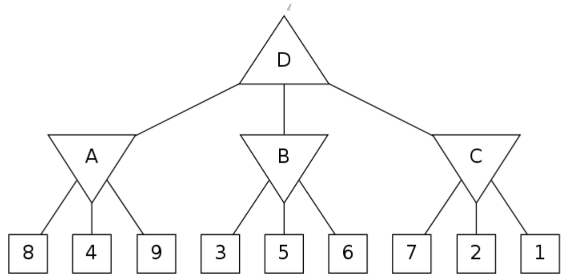



Figure 1: game tree

Consider the game tree shown in Figure 1. Triangles that point up, such as at the top node (root), represent choices for the maximizing player; triangles that point down represent choices for the minimizing player. Assuming both players act optimally, use alpha-beta pruning to find the value of the root node. The search goes from left to right; when choosing which child to visit first, choose the left-most unvisited child. In the blank spaces below, enter the values of the labeled nodes. Then, select the leaf nodes that don't get visited due to pruning.

Hint: Note that the value of a node where pruning occurs is not necessarily the maximum or minimum (depending on which node) of its children. When you prune on conditions $V > \beta$ or $V < \alpha$, assume that the value of the node is $V$.

A: ***4*** B: ***3*** C: ***2*** D: ***4***

Check the boxes next to leaf nodes that are not visited due to pruning.

☐ 8 ☐ 4 ☐ 9 ☐ 3 √ ***5*** √ ***6*** ☐ 7 ☐ 2 √ ***1***

Explanation. A: $\alpha = -\infty$ and $\beta = \infty$. Because $\alpha = -\infty$, there will be no pruning. Intuitively, this means that any value that the minimizer finds might be used by the maximizer.

B: $\alpha = 4$ and $\beta = \infty$. The first leaf value, 3, is less than 4, so the remaining children can be pruned and B gets value 3. Intuitively, B will never take a value greater than 3, and D will never select a value less than 4, so you know that D will never select the middle action.

C: $\alpha = 4$ and $\beta = \infty$. The first leaf value, 7, is greater than 4, so the other children must be checked. The second leaf value, 2, is less than 4, so the remaining child can be pruned and C gets value 2.

D: Because $\beta = \infty$, none of D's children can be pruned, which is always true for the root. D then takes the max of its children, and has value 4 (which, because none of its children were pruned, is identical to its value when running minimax without pruning)

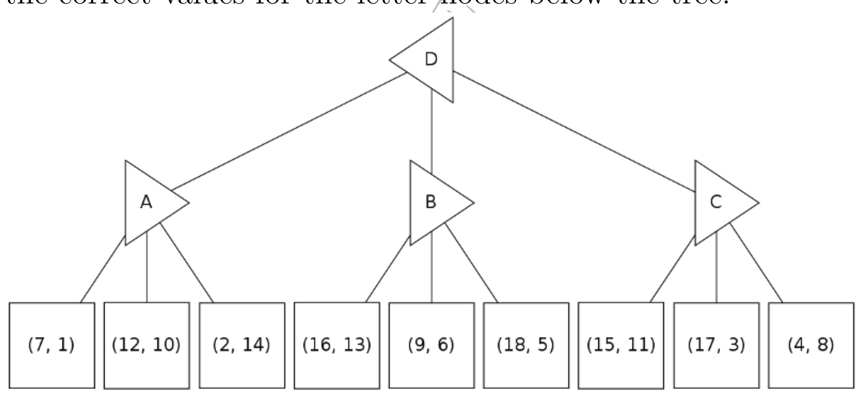
5. (8 points) NON-ZERO-SUM GAMES

   (a) The standard minimax algorithm calculates worst-case values in a zero-sum two player game, i.e. a game for which in all terminal states $s$, the utilities for players A (MAX) and B (MIN) obey $U_A(s) + U_B(s) = 0$. In this zero-sum setting, we know that $U_A(s) = -U_B(s)$, so we can think of player B as simply minimizing $U_A$.

   In this problem, you will consider the non-zero-sum generalization, in which the sum of the two players' utilities are not necessarily zero. The leaf utilities are now written as pairs $(U_A, U_B)$. In this generalized setting, A seeks to maximize $U_A$, the first component, while B seeks to maximize $U_B$, the second component.

   Consider the non-zero-sum game tree below. Note that left-pointing triangles (such as the root of the tree) correspond to player A, who maximizes the first component of the utility pair, whereas right-pointing triangles (nodes on the second layer) correspond to player B, who maximizes the second component of the utility pair.

   Propagate the terminal utility pairs up the tree using the appropriate generalization of the minimax algorithm on this game tree. In case of ties, choose the leftmost child. Select the correct values for the letter nodes below the tree.




   Your answer should be in the format (X,Y), where X is the value of Player A and Y is the value of Player B at a node (for example "(7,1)"). Note that the answer check is sensitive to formatting, so be sure to include the parenthesis and do not use any spaces in your answers.

   A: ___*(2,14)*___ B: ___*(16,13)*___ C: ___*(15,11)*___ D: ___*(16,13)*___

Explanation. Each player only considers their own utility.

A gets (2, 14) because that is its choice with the highest utility on the right.

B gets (16, 13) because that is its choice with the highest utility on the right.

C gets (15, 11) because that is its choice with the highest utility on the right.

D gets (16, 13) because that is its choice with the highest utility on the left.

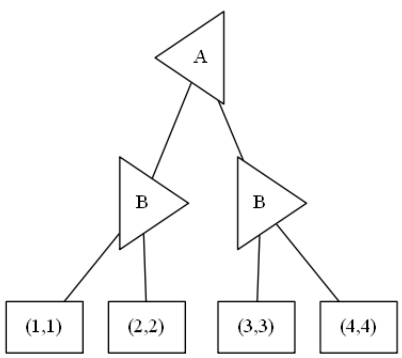
(b) In this problem, you will again consider the non-zero-sum generalization, in which the sum of the two players' utilities are not necessarily zero. The leaf utilities are now written as pairs $(U_A, U_B)$. In this generalized setting, A seeks to maximize $U_A$, the first component, while B seeks to maximize $U_B$, the second component.

Assume that your generalization of the minimax algorithm calculates a value $(U_A^*, U_B^*)$ for the root of the tree. Assume no utility value for A or for B appears more than once in the terminal nodes (this means there will be no need for tie-breaking). Which of the following statements are true?

- √ ***Assuming A and B both play optimally, player A's outcome is guaranteed to be exactly $U_A^*$.***
- √ ***Assuming A and B both play optimally, player B's outcome is guaranteed to be exactly $U_B^*$.***
- ☐ Assuming B plays sub-optimally (but A plays optimally), A's outcome is guaranteed to be at least $U_A^*$.

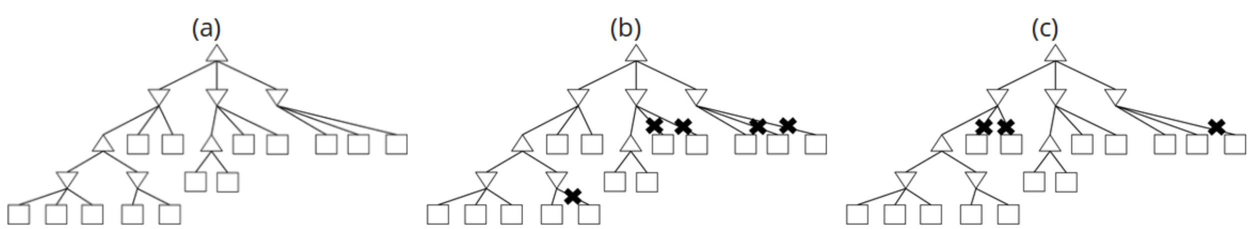
Explanation. The generalized algorithm finds the value assuming that each player plays optimally. However, because of the non-zero sum settings, sub-optimal play by one player does not necessarily improve the result for the other player.

Consider the following case where all utilities are the same for both players.




If both players play optimally, they would both end with a score of 4. However, if B plays sub-optimally, they both end up with a score of 3.

6. (10 points) (a) Assume we run $\alpha$ - $\beta$ pruning, expanding successors from left to right, on a game with tree as shown in Figure (a) below.




Which of the following statements are true?

- √ ***There exists an assignment of utilities to the terminal nodes such that no pruning will be achieved (shown in Figure (a)).***
- √ ***There exists an assignment of utilities to the terminal nodes such that the pruning shown in Figure (b) will be achieved.***
- ☐ There exists an assignment of utilities to the terminal nodes such that the pruning shown in Figure (c) will be achieved.
- ☐ None of the above.

Explanation. Perhaps the simplest check is as follows: pruning of children of a minimizer node $m$ is possible (for some assignment to the terminal nodes), when both of the following conditions are met: (i) the value of another child of $m$ has already been determined, (ii) somewhere on the path from $m$ to the root node, there is a maximizer node $M$ for which an alternative option has already been explored. The pruning will then happen if any such alternative option for the maximizer had a higher value than the value of the "another child" of $m$ for which the value was already determined.

a) One such assignment (starting with shallowest leaves):

3,3 — 5,5 — 6,6,6

4,4

1,1,1 — 2,2

b) One such assignment:

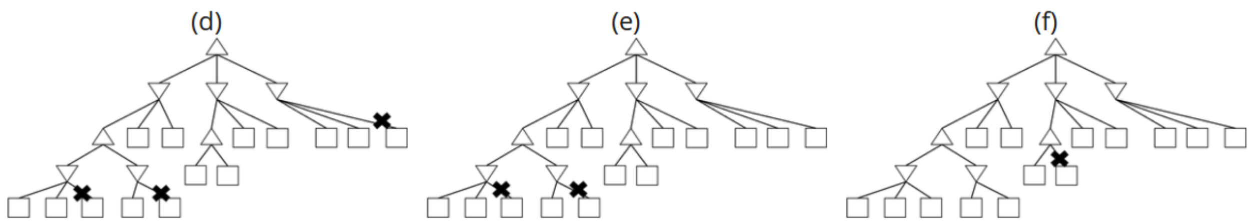
2,2 — x,x — 0,x,x

0,0

1,1,1 — 0,x

c) The left-most child of the root would have $\alpha = -\infty$, so its direct children can never be pruned because they will always be greater than $-\infty$. Intuitively the root hasn't seen anything else yet, so any value returned by the minimizer might end up being taken by its parent.

(b)




- ☐ There exists an assignment of utilities to the terminal nodes such that the pruning shown in Figure (d) will be achieved.
- ☐ There exists an assignment of utilities to the terminal nodes such that the pruning shown in Figure (e) will be achieved.
- ☐ There exists an assignment of utilities to the terminal nodes such that the pruning shown in Figure (f) will be achieved.
- √ ***None of the above.***

Explanation. d,e) The left-most node has $\alpha = -\infty$, so any value seen will be greater than $\alpha$, and thus none of its children can be pruned.

f) The leaf that we want to prune has $\beta = \infty$, so any value in the first child will be less than $\beta$, and thus no pruning is possible

7. (9 points) Suboptimal Strategies

   (a) Player MAX and player MIN are playing a zero-sum game with a finite number of possible moves. MAX calculates the minimax value of the root to be $M$. You may assume that at every turn, each player has at least 2 possible actions. You may also assume that a different sequence of moves will always lead to a different score (i.e., no two terminal nodes have the same score). Which of the following statements are true?

   - √ ***Assume MIN is playing sub-optimally at every turn, but MAX does not know this. The outcome of the game could be larger than $M$ (i.e. at least as good a MAX).***
   - ☐ Assume MIN is playing sub-optimally at every turn. If MAX plays according to the minimax strategy, the outcome of the game could be less than $M$.

Explanation. a) Consider a toy example in which there is only one turn, MAX has two actions, and MIN has two actions. For both of MAX's actions, one terminal node is positive, and the other is negative. M would be one of the negative values, while the actual outcome would be one of the positive values.

b) If MAX is playing optimally, the minimax value is only a lower bound on the outcome of the game. If MIN plays suboptimally, that can only improve MAX's options each turn and thus improve the outcome for MAX.

(b) For this question, assume that MIN is playing randomly (with a uniform distribution) at every turn, and MAX knows this. Which of the following statements are true?

- ☐ There exists a policy for MAX such that MAX can guarantee a better outcome than $M$.
- √ ***There exists a policy for MAX such that MAX's expected outcome is better than*** $M$***.***
- ☐ To maximize his or her expected outcome, MAX should play according to the minimax strategy (i.e. the strategy that assumes MIN is playing optimally).

Explanation. a) MIN could randomly choose the optimal action at every turn, which would result in an outcome of M if max plays according to the minimax policy, or possibly less if max plays according to a different policy.

b) The policy found using Expectimax will have a higher expected outcome than M.

c) Expectimax gives MAX a higher expected outcome than the minimax strategy.

(c) Which of the following statements are true?

- √ ***Assume MIN is playing sub-optimally at every turn. MAX following the minimax policy will guarantee an outcome that is better than or equal to*** $M$***.***
- √ ***Assume MIN is playing sub-optimally at every turn, and MAX knows exactly how MIN will play. There exists a policy for MAX to guarantee an outcome that is better than or equal to*** $M$***.***

Explanation. a) MIN playing sub-optimally every turn will only possibly improve MAX's choices each turn, because the minimax policy assumes that MIN will choose the worst possible move from MAX's perspective. Thus, at each step, MAX is guaranteed to have a value that is better than or equal to the minimax value.

b) As stated above, the minimax policy will guarantee MAX an outcome that is better than or equal to M.

8. (9 points) RATIONALITY OF UTILITIES

(a) Consider a lottery L = [0.2 , A; 0.3, B; 0.4, C; 0.1, D], where the utility values of each of the outcomes are U(A) = 1, U(B) = 3, U(C) = 5, U(D) = 2. What is the utility of this lottery, U(L)?

*Answer:* ***3.3***

Explanation. U(L) = 0.2 U(A) + 0.3 U(B) + 0.4 U(C) + 0.1 U(D) = 3.3.

(b) Consider a lottery L1 = [0.5, A; 0.5, L2], where U(A) = 4, and L2 = [ 0.5, X; 0.5, Y] is a lottery, and U(X) = 4, U(Y) = 8. What is the utility, U(L1), of the the first lottery?

*Answer.* **5**

Explanation. U(L2) = 0.5 U(X) + 0.5 U(Y) = 2 + 4 = 6
U(L1) = 0.5 U(A) + 0.5 U(L2) = 2 + 3 = 5

(c) Assume $A \succ B$, $B \succ L$, where $L = [0.5, C; 0.5, D]$, and $D \succ A$. Assuming rational preferences, which of the following statements are guaranteed to be true?

$\surd$ $A \succ L$ $\surd$ $A \succ C$ □ $A \succ D$ $\surd$ $B \succ C$ □ $B \succ D$

Explanation. a) $A \succ B \succ L$, so by transitivity $A \succ L$

b) $A \succ L \implies A \succ D \vee A \succ C$. Because $D \succ A$, then $A \succ C$ must be true.

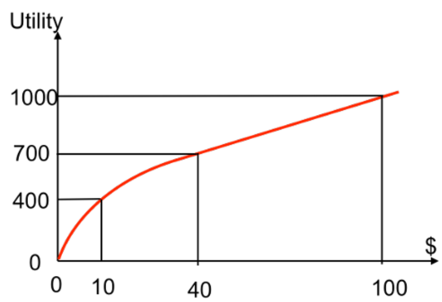
c) $D \succ A$ means this is false.

d) $D \succ A \succ B \implies D \succ B$, so for the same reasoning as (b) this is true

e) $D \succ A \succ B \implies D \succ B$, means this is false.

9. (6 points) Certainty Equivalent Values

Consider the utility function shown below.




Under the above utility function, what is the certainty equivalent monetary value in dollars (\$) of the lottery [0.6, \$0 ; 0.4, \$100]?

I.e., what is $X$ such that $U(\$X) = U([0.6, \$0; 0.4, \$100])$?

[*Hint.* Keep in mind that $U([p, A; 1-p, B])$ is not equal to $U(pA + (1-p)B)$.]

*Answer.* **10**

Explanation. $U([0.6, \$0; 0.4, \$100]) = 0.6U(\$0) + 0.4U(\$100) = 400$

$U(\$10) = 400$

10. (14 points) Preferences and Utilities

Our Pacman board now has food pellets of 3 different sizes - pellet $P_1$of radius 1, $P_2$ of radius 2 and $P_3$ of radius 3. In different moods, Pacman has different preferences among these pellets. In each of the following questions, you are given Pacman's preference for the different pellets. From among the options pick the utility functions that are consistent with Pacman's preferences, where each utility function $U(r)$ is given as a function of the pellet radius $r$, and is defined over non-negative values of $r$.

(a) $P_1 \sim P_2 \sim P_3$

$\surd$ $U(r) = 0$ $\surd$ $U(r) = 3$ □ $U(r) = r$ □ $U(r) = 2r{+}4$ □ $U(r) = -r$
□ $U(r) = r^2$ □ $U(r) = -r^2$ □ $U(r) = \sqrt{r}$ □ $U(r) = -\sqrt{r}$
□ Irrational preferences!

Explanation. Because all three sizes are preferred equally, $U(r)$ has to return the same value for $r = 1, 2, 3$. The only functions that do so are those that do not depend on $r$.

(b) $P_1 \prec P_2 \prec P_3$

☐ $U(r) = 0$ ☐ $U(r) = 3$ $\surd$ $U(r) = r$ $\surd$ $U(r) = 2r + 4$ ☐ $U(r) = -r$
$\surd$ $U(r) = r^2$ ☐ $U(r) = -r^2$ $\surd$ $U(r) = \sqrt{r}$ ☐ $U(r) = -\sqrt{r}$
☐ Irrational preferences!

Explanation. Higher radii are preferred over lower ones, so increasing functions of $r$ satisfy the constraints.

(c) $P_1 \succ P_2 \succ P_3$

☐ $U(r) = 0$ ☐ $U(r) = 3$ ☐ $U(r) = r$ ☐ $U(r) = 2r + 4$ $\surd$ $U(r) = -r$
☐ $U(r) = r^2$ $\surd$ $U(r) = -r^2$ ☐ $U(r) = \sqrt{r}$ $\surd$ $U(r) = -\sqrt{r}$
☐ Irrational preferences!

Explanation. Lower radii are preferred over higher ones, so decreasing functions of $r$ satisfy the constraints.

(d) $(P_1 \prec P_2 \prec P_3)$ and $(P_2 \prec ($ 50-50 lottery among $P_1$ and $P_3))$

☐ $U(r) = 0$ ☐ $U(r) = 3$ ☐ $U(r) = r$ ☐ $U(r) = 2r + 4$ ☐ $U(r) = -r$
$\surd$ $U(r) = r^2$ ☐ $U(r) = -r^2$ ☐ $U(r) = \sqrt{r}$ ☐ $U(r) = -\sqrt{r}$
☐ Irrational preferences!

Explanation. The first constraint means that $U(r)$ must be increasing, and the second constraint means that the rate at which it is increasing must be increasing as well, and $r^2$ is the only function that is of the ones provided.

(e) $P_1 \succ P_2 \succ P_3$ and $P_2 \succ ($ 50-50 lottery among $P_1$ and $P_3)$

☐ $U(r) = 0$ ☐ $U(r) = 3$ ☐ $U(r) = r$ ☐ $U(r) = 2r + 4$ ☐ $U(r) = -r$
☐ $U(r) = r^2$ $\surd$ $U(r) = -r^2$ ☐ $U(r) = \sqrt{r}$ ☐ $U(r) = -\sqrt{r}$
☐ Irrational preferences!

Explanation. The first constraint means that $U(r)$ must be decreasing, and the second constraint means that $U(2) > 0.5 * U(1) + 0.5 * U(3)$.

(f) $P_1 \prec P_2 \prec P_3$ and (50-50 lottery among $P_2$ & $P_3$) $\prec$ (50-50 lottery among $P_1$ & $P_2$)

☐ $U(r) = 0$ ☐ $U(r) = 3$ ☐ $U(r) = r$ ☐ $U(r) = 2r + 4$ ☐ $U(r) = -r$
☐ $U(r) = r^2$ ☐ $U(r) = -r^2$ ☐ $U(r) = \sqrt{r}$ ☐ $U(r) = -\sqrt{r}$
$\surd$ ***Irrational preferences!***

Explanation. $P_3 \succ P_1$ by transitivity, and since both lotteries have equal chances for $P_2$, it can be ignored when comparing the two. So the last constraint is essentially $P_3 \prec P_1$, which makes the preferences irrational.

(g) Which of the following would be a utility function for a risk-seeking preference? That is, for which utility(s) would Pacman prefer entering a lottery for a random food pellet, with expected size $s$, over receiving a pellet of size $s$?

☐ $U(r) = 0$ ☐ $U(r) = 3$ ☐ $U(r) = r$ ☐ $U(r) = 2r + 4$ ☐ $U(r) = -r$
$\surd$ $U(r) = r^2$ ☐ $U(r) = -r^2$ ☐ $U(r) = \sqrt{r}$ $\surd$ $U(r) = -\sqrt{r}$

Explanation. Functions that are either decreasing slower than linearly, like $-\sqrt{r}$, or increasing faster than linearly, like $r^2$, satisfy this.