National University of Singapore
School of Computing

Semester 2, AY2023-24

CS4246/CS5446 AI Planning and Decision Making

Tutorial Week 6: DT, DA and MDP

Guidelines

- You can discuss the content of the questions with your classmates.
- However, everyone should work on and be ready to present ALL the solutions.
- Your attendance is marked in the tutorial and participation noted to award class participation marks.

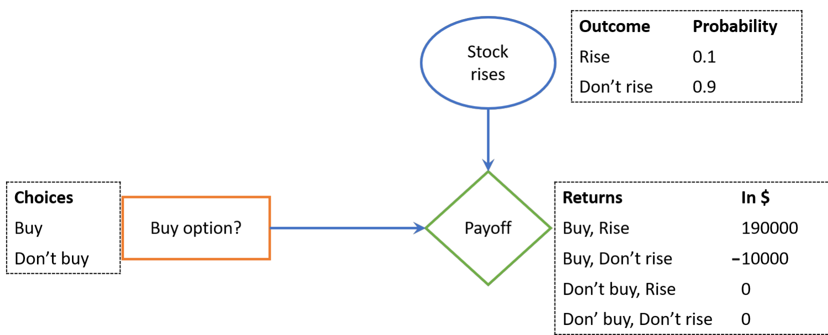
Problem 1: Basic Risky Decision

Note: This question is brought forward from the last week's tutorial to give you the context for the next problem

Richie Bean is trying to strike it big in the stock market during the economic downturn. He is considering buying some options to a very risky stock on a diamond mine in Africa. There is only a 10% chance that the stock price will rise if he exercises his options, but the payoff is $200,000. It costs $10,000 to buy and exercise the options. The alternative is not to buy at all, in which case Mr. Bean's profit is zero.

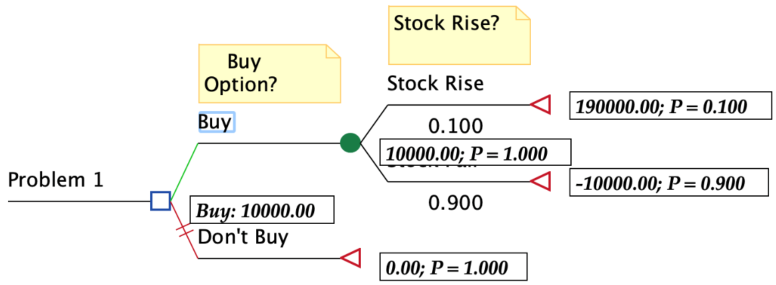
a. Draw an influence diagram to represent Mr. Bean's problem. Clearly indicate all the options/outcomes and numbers. Should he buy the options? Use the solution approaches mentioned in the lecture to substantiate your answer.

Solution:



b. Draw an decision tree to represent Mr. Bean's problem. Clearly indicate all the options/outcomes and numbers. Should he buy the options? Show all the details in your decision tree.

Solution:



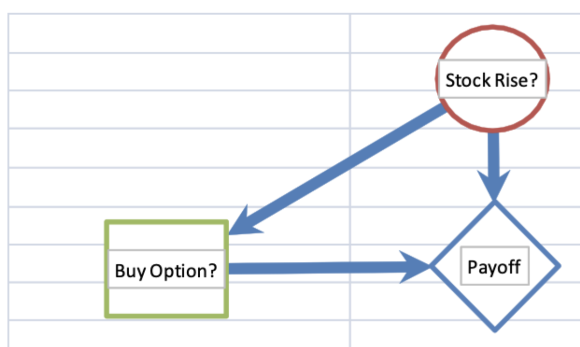
EMV for buy \$10000, don't buy \$0.

Problem 2: Value of Perfect Information

a. Represent the hypothetical situation where Mr. Bean will get perfect information before he makes the decision. How to represent this situation in an influence diagram? Clearly indicate all the options/outcomes and numbers.

Solution:

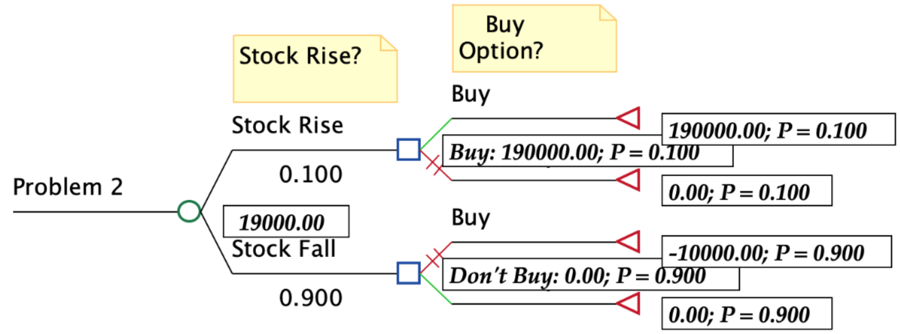
[Solution]: (Numbers should be shown as well, but are omitted here)



b. How to represent this situation in the decision tree? What is the expected value of the decision with perfect information?

Solution:

Solution:



EMV with perfect information is $19,000
EVPI = $19,000 - $10,000 = $9,000 ---- Max amount to pay for Perfect information

Problem 3: Scrooge and his nephews

Uncle Scrooge is deciding which of his three grand-nephews, Huey, Duey, and Louie, is going to inherit his wealth and business empire. To help him in his assessment, Uncle Scrooge sets up the following lottery, where the choice is between:
Option (A): Receiving $4900 (with probability 1), OR
Option (B): participate in a lottery which has a 60% probability of winning $7000 and a 40% probability of winning $20.

a) Huey's utility function for wealth (money) x is represented by an exponential utility function: $U(x) = log_{10}(x)$, for all $x > 0$. Assuming that Huey would choose the option with the higher utility, which option would he choose?
Solution:
Huey would choose Option (A).
Option (A) utility = 3.69 (by plugging in the values)
Option (B) utility = 2.82.

b) Duey's utility function for wealth x is represented by $U(x) = \sqrt{x}$, for all $x > 0$. Assuming that Duey would choose the option with the higher utility, which option would he choose?
Solution:
Duey would choose Option (A).
Option (A) utility = 70
Option (B) utility = 51.98

c) Louie's utility function for wealth $x, \forall x > 0$ is represented by the following utility function

$$U(x) = \begin{cases} x, & 0 \leq x \leq 5000 \\ (10000 - x), & 5000 < x \leq 10000 \\ 10000, & x > 10000 \end{cases}$$

Assuming Louie would choose the option with the higher utility, which option would Louie choose?

Solution:

Louie would choose Option (A).
Option (A) utility = 4900
Option (B) utility = 1808

d) Uncle Scrooge Scrooge needs to choose a successor who is rational and reasonably risk averse, so that he can preserve and yet grow the business. After fully understanding each of their utility functions, which of Huey, Duey, and Luey would Uncle Scrooge consider to be his successor? There could be more than one choice.

Solution:

Risk averse means that the utility function is concave. When the second derivate of the function is negative, it is a concave function.

Both Huey and Duey's utility functions are rational and risk averse. However, Louie's utility function does not conform to monotonicity criteria, making him irrational.

Uncle Scrooge would select either Huey or Duey as his successor.

Problem 4: Formulating Markov Decision Processes

Specify the following problems as a Markov decision process, *i.e.* specify the state space, the actions, the transition functions, and the reward function. What is the (approximate) size of the state space and the action space?

- The traveling salesman problem. A salesman must visit every city in a graph and minimize travel time and is constrained not to visit any city twice.

  Solution:

  Assume that there are N cities. The state can be specified by a pair (V, c) where V is a subset of cities that have already been visited and c is the city that the salesman is currently at. V can take 2^N possible values and c can take N values, so the state space size is $O(N2^N)$. The actions are to move from the current city to another city that is connected to the current city. There is a special initial state where the salesman starts with none of the cities visited, and a special terminal state where all the cities have been visited and the salesman is back at the starting city. The action space size is N. Transition is deterministic: moving from c to c' changes (V, c) to $(V \cup c, c')$. At the terminal state, all actions self-loop with zero reward. Rewards for actions at other states is the negative of the edge weight of the edge (c, c') for moving from c to c' if c' has not been visited and $-\infty$ (or a very large negative number) otherwise. This is a finite horizon MDP with horizon N. It is also deterministic, so can be solved using deterministic planning methods (although can also be represented as a MDP).

- Inventory control. The company has space to store N items. At the end of each day, the company will make an order to increase the number of items up to $M \leq N$. Placing an order cost c for each time an order is made. If there is not enough items in the inventory to meet the orders for the day, a back order has to be made at the cost of b per unit back ordered (up to a known maximum of B units). There is a holding cost of 1 for each item in the inventory at the end of the day.

Solution:
The state space is the set of integers ranging from $-B$ to N indicating how many items is in stock where a negative number indicates the number that needs to be backordered. The state space size is $N + B + 1$. The actions are to order items to increase the number to M for $M = 0, \ldots, N$, with $N + 1$ actions. The transition is deterministic and transitions the system to state M. This is followed by subtracting the day's demand d, where d is drawn from the demand probability $\Pr(d)$. For state $s \geq 0$, the action to increase the number to M has reward $-c - s$ unless no order is required giving reward -s. For state $s < 0$, the action to increase the number to M has reward $sb - c$ (s is negative so sb is negative), unless no order is required giving reward sb.