National University of Singapore
School of Computing

Semester 1, AY2023-24

CS4246/CS5446 AI Planning and Decision Making

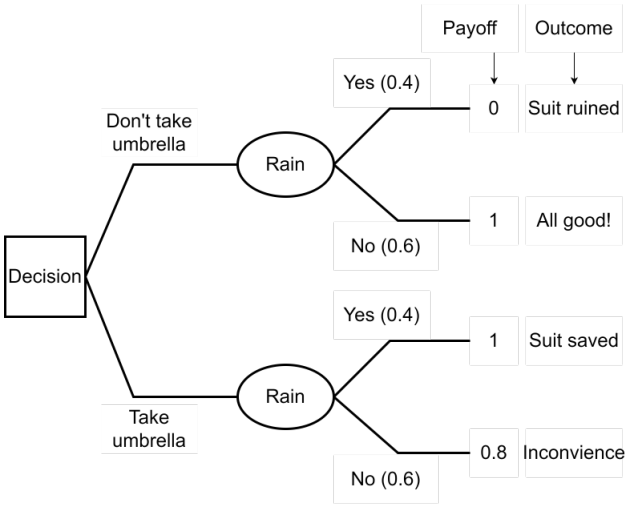
Tutorial Week 7: DA and MDP

Guidelines

You may discuss the content of the questions with your classmates. But everyone should work on and be ready to present ALL the solutions.

Problem 1: VPI

Mr. Huges has bought a new suit that he wants to wear to office. However, it looks like it might rain. Mr. Huges has a long walk to the MRT. If it rains his suit will be ruined. If it doesn't rain, no harm done. His umbrella will protect his suit, but he hates the inconvenience of carying it around all day. If it rains and he has the umbrella the suit is saved. There is a 40% chance that it will rain. His decision problem is represented as the following decision tree.



He can call AccurateWeather, a perfect information service that tells if it rains or not perfectly.

1. Draw the decision tree to include this perfect information source in the decision making.

2. What is the "value" (in terms of the payoff) that Mr. Huges should pay to AccurateWeather to gather information about the rain?

Problem 2: 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.
- 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.

Problem 3: Value Iteration

Consider the following 2 state, 2 action MDP with discount factor 0.9.

$P(s_1\|s_1,a_1)$	$P(s_2\|s_1,a_1)$	$P(s_1\|s_2,a_1)$	$P(s_2\|s_2,a_1)$
0.9	0.1	0	1

$P(s_1\|s_1,a_2)$	$P(s_2\|s_1,a_2)$	$P(s_1\|s_2,a_2)$	$P(s_2\|s_2,a_2)$
0.1	0.9	0	1

$R(s_1,a_1)$	$R(s_1,a_2)$	$R(s_2,a_1)$	$R(s_2,a_2)$
1	0	3	3

1. Assume a finite horizon problem with horizon 1 (only 1 action is to be taken). What is the utility or value function and the optimal action in each state?
2. Assume a finite horizon problem with horizon 2 (2 actions to be taken). What is the utility or value function and the optimal action in each state?
3. What is the optimal infinite horizon policy?