Assignment 4
TDT4171 Artificial Intelligence Methods

ALEKSANDER ELVEBAKK

March 10, 2020

Contents

1 Utility

a)

In this task I will assume we are talking about epleslang given the useless description of what we are talking about here. Gabriel's slope is increasing for the number of apples, that is how long he is there, so he is risk seeking. Gustav and Sonja has a linear utility function, so they are risk neutral. They do not care how long they are staying for apples, they just count the apples as utility. Maria's slope is decreasing and she is thus risk-averse.

b)

In the risk-seeking lottery you pay 10, and you have 1% chance of winning 2000 kroner. That is formulated as follows:

$$L_a = [(0.01, 1990), (0.99, -10)]$$

The expected monetary value is then

$$EMV = 0.01 * 1990 + 0.99 * -10 = 10$$

In the risk-averse lottery, you get payed 11.

$$L_b = [(1, 11), (0, 0)]$$

Here the expected monetary value is obviously 11
Using the utility function $U = x^3$ and the following defintion of expected utility of lottery:

$$EU(L) = p_1 U(r_1) + p_2 U(r_2)$$

Utility of expected monetary value:

$$U(E(L)) = (E(L))^3$$

	Risk-seeking	Risk-averse
Lottery	[(0.01, 1990), (0.99, -10)]	[(1, 11),(0, 0)]
Expected utility of lottery	78805000	729
Utility of expected monetary value	1000	1331

2 Decision Network

a)

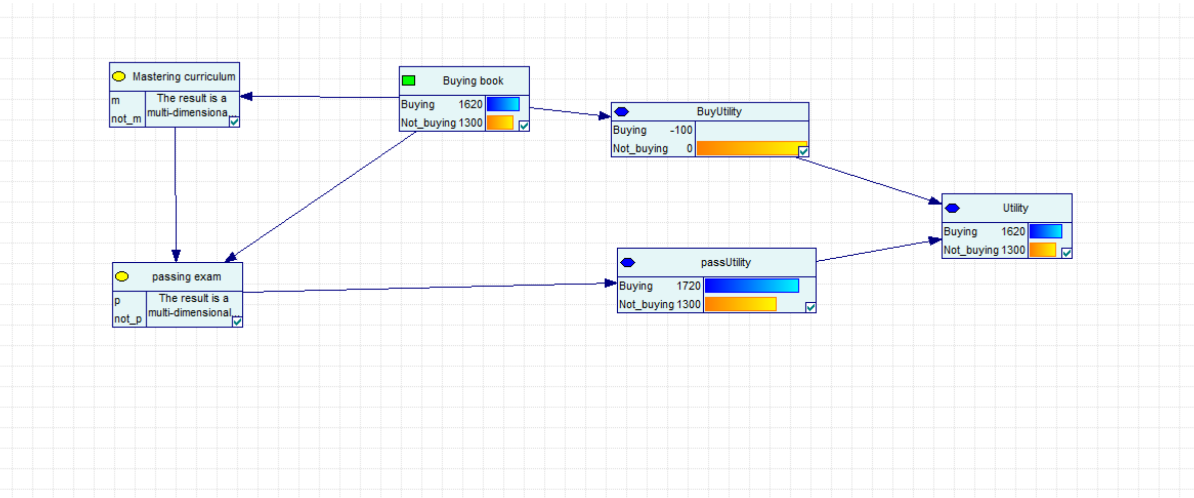

Figure 1: Network in task 2a)

b)

Buying:

$$
\begin{aligned}
E(U(b)) &= E(U_1(b)) + E(U_2(b)) \\
E(U_1(b)) &= -100 \\
E(U_2(b)) &= P(p|b)2000 \\
P(p|b) &= P(p|\neg m, b)P(\neg m|b) + P(p|m, b)P(m|b) \\
P(\neg m|b) &= 1 - P(m|b) = 0.1 \\
E(U(b)) &= 1620
\end{aligned}
$$

Not Buying:

$$
\begin{aligned}
E(U(\neg b)) &= E(U_1(\neg b)) + E(U_2(\neg b)) \\
E(U_1(\neg b)) &= 0 \\
E(U_2(\neg b)) &= P(p|\neg b)2000 \\
P(p|\neg b) &= P(p|\neg m, \neg b)P(\neg m|\neg b) + P(p|m, \neg b)P(m|\neg b) \\
P(\neg m|\neg b) &= 1 - P(m|\neg b) = 0.3 \\
E(U(\neg b)) &= 1300
\end{aligned}
$$

c)

A rational agent should always choose the alternative which maximizes the utility function, so he should be buying the book.

3 Markov decision process

a)

This iteration will use the Bellman update

$$U_{i+1}(s) \leftarrow R(s) + \gamma \max_{a \in A(s)} \sum_{s'} P(s'|s,a)U_i(s')$$

For this task, this will be (here using state 1 as an example)

$$\begin{aligned}
U_{i+1}(1) &= Rewrd(1) + \gamma \max(L(1), R(1)) \\
L(1) &= P_L(1 \to 1)U_i(1) + P_L(1 \to 2)U_i(2) + P_L(1 \to 3)U_i(3), \\
R(1) &= P_R(1 \to 1)U_i(1) + P_R(1 \to 2)U_i(2) + P_R(1 \to 3)U_i(3)
\end{aligned}$$

The same logic goes for L(2), L(3), R(3) and R(3) This gives the following values in iteration 1 and 2

Initial	U[1]=0	U[2]=0	U[3] = 0
Iteration 1	U[1]=0	U[2]=1.0	U[3] = 0
Iteration 2	U[1]=0.375	U[2]=1	U[3] = 0.375

b)

$$\begin{aligned}
& L = P_L(1 \to 1)U(1) + P_L(1 \to 2)U(2) + P_L(1 \to 3)U(3) \\
\Rightarrow\ & L = 0.6875 \\
& R = P_R(1 \to 1)U(1) + P_R(1 \to 2)U(2) + P_R(1 \to 3)U(3) \\
\Rightarrow\ & R = 1.0625
\end{aligned}$$

4 Value Iteration

b)

	1	2	3	4
1	1.13	0.48	1.46	0.16
2	1.51	-10.0	2.08	-10.0
3	3.3	5.78	5.53	-10.0
4	-10.0	7.07	8.35	10.0

Figure 2: Task 4b) utilities of tiles

c)

	1	2	3	4
1	down	up	down	up
2	down		down	
3	right	down	down	
4		right	right	right

Figure 3: Task 4c) optimal actions in each state