Computer Science 182

Jaemin Cheun, HUID:40871209

Assignment 5

Part B

(1) MDP is a 4-tuple (not including the discount factor) (S, A, R, T):

(i) S is a finite set of states.
The states are given as: $S = \{\text{cool, warm, off}\}$, which we will simplify as $S = \{\text{C, W, O}\}$

(ii) A is a finite set of actions.
The actions are given as: $A = \{\text{fast, slow}\}$, which we will simplify as $A = \{\text{F, S}\}$

(iii) R is the reward model:

State	Action	Reward
C	F	10
C	S	4
W	F	10
W	S	4
O	F	0
O	S	0

(iv) Transition model T^a for each action a:

$$T^F = \begin{bmatrix} 1/4 & 3/4 & 0 \\ 0 & 7/8 & 1/8 \\ 0 & 0 & 1 \end{bmatrix}, T^S = \begin{bmatrix} 1 & 0 & 0 \\ 1/4 & 3/4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where columns represent target state (C,W,O) and rows represent source state (C,W,O)

(2)

k	$\pi_k(C)$	$\pi_k(W)$	$\pi_k(O)$	$V_k(C)$	$V_k(W)$	$V_k(O)$
0	F	F	F	0	0	0

Problem 5

1. HTN:
How to build "sample collection-site booths" will be best modeled by a HTN. This planning problem can be decomposed into smaller compound tasks and primitive tasks. In this case, there is no obvious uncertainty involved, so MDP is not required.

2. MDP: