CS499

Assignment 1

Sept 11, 2006

1. The assignment is 20 points total.

2. It will be due back on or before Sept 18, 2006 at 2 pm. Please turn in a hard copy of your solutions either in class or in our office hours.

3. This assignment must be done individually. Please do not discuss either the questions or answers with others.

4. Please feel free to send questions to tambe@usc.edu or bowring@usc.edu and we will respond to the class mailing list.
Question 1: For each of the following sentences, write the BDI logical equivalent, using the box and diamond notation. (3 pt)

A. Powell thinks Speedy believes that eventually pool 1 will be safe.

B. Powell thinks Speedy wrongly believes that eventually pool 1 will be safe.

C. Speedy always believes that pool 1 is unsafe.

Question 2: Are the following statements true or false? Explain in one or two sentences why. (3 pt)

A. In Cohen and Levesque logic, it is inconsistent for agent x to believe both (Bel x \( \Diamond P \)) and (Bel x \( \Diamond \neg P \))

   TRUE ----   FALSE ----

   Why?

B. In Cohen and Levesque logic, it is consistent for agent x to believe both (Bel x \( \Diamond P \)) and (Bel x \( \neg P \))

   TRUE ----   FALSE ----

   Why?

C. In Cohen and Levesque logic it is possible for an agent to have a belief (Bel x \( \neg \Diamond P \)) and a goal (Goal x \( \neg ( \neg P ) \))

   TRUE ----   FALSE ----

   Why?
Question 3: (3 pt)
*From Runaround:* When Powell and Donovan managed to get Oxalic acid thrown near Speedy, Speedy walked away from the Selenium pool (even if temporarily). Write down Speedy’s PGOAL when it was walking way from the Selenium pool.
*Hint: There may not be one perfect answer.*

Question 4: (4 pt)
A. The Enterprise is starting up an intra-ship lottery. Tickets are going to cost $2.50 and the odds of winning will be 1/50. However, there is some disagreement as to how much prize money should be offered; the potential prizes are: $120, $125, $130, $150. For each of the following crew members, list which prize(s) would induce him to play the lottery and briefly explain why. Please assume whatever degree of risk aversion or risk seeking that will allow you to include the maximum number of possible prizes in your answer.
(i) Commander Data (Risk Neutral)
(ii) Captain Picard (Risk Averse)
(iii) Commander Reiker (Risk Seeking)

B. Does the following graph show the utility curve of a risk-seeking, risk-averse or risk-neutral person? Please briefly justify your answer.

![Utility Curve Graph](image-url)
Question 5: (4 pt)

- States: \{s1, s2, s3, s4, s5, s6, s7\}; s1 is the start state

- Actions: In states s1, s2 and s3, two actions possible: LEFT and RIGHT
  In state s4, s5, s6, s7 no action possible. These are terminal.

- Reward: $R(s1) = R(s2) = R(s3) = 0$
  $R(s4) = -1$
  $R(s5) = 1$
  $R(s6) = 2$
  $R(s7) = -1$

- Transition probabilities

  **ACTION: LEFT**
  \[
  \begin{array}{cccccccc}
  & s1 & s2 & s3 & s4 & s5 & s6 & s7 \\
  s1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
  s2 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \\
  s3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
  \end{array}
  \]

  **ACTION: RIGHT**
  \[
  \begin{array}{cccccccc}
  & s1 & s2 & s3 & s4 & s5 & s6 & s7 \\
  s1 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\
  s2 & 0 & 0 & 0 & 0.75 & 0.25 & 0 & 0 \\
  s3 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\
  \end{array}
  \]

Given the above MDP please answer the following questions. In answering the questions, please provide a brief explanation of how you got the answers:

A. After one step of value iteration in state S2, what is the utility $U(S2)$?
B. After one step of value iteration in state S3, what is the utility $U(S3)$?
C. After two steps of value iteration in state S1, what is the utility $U(S1)$?
D. Now write down the optimal policy for this MDP.

*Hint: It may help you in terms of answering this question to think of the state space (states and transitions) as was shown in class.*
Question 6: (3 pt)
For each of the following MDPs, draw the optimal policy ($\pi^*$) and give an intuitive answer as to why it is optimal (note, you are not required to run value iteration). For each subquestion, $A = \{\text{North, South, East, West}\}$ and the transition function $P$ is the same in every square. For a given action, with probability 0.8, the action will succeed and with a probability of 0.1 it will be deflected into an action at a 90 degree angle from the intended action. For example, if the action taken were South, $P = 0.8$ that the agent will move south, $P = 0.1$ that the agent will move west and $P = 0.1$ that the agent will move east.

a) $R(I) = +0.5$

```
+10
-10
```

b) $R(I) = -0.5$

```
-100
-10
```

C. $R(I) = 0$

```
-\infty  +50
```
