CISC 7412 Fall 2011, Homework 2

1. Figure 1 (next page) shows a an agent situated in an enviroment. Using the notation from Lecture #2, we can think of the environment E being made up of 36 states:

   $$e_{0,0}, e_{0,1}, \ldots$$

   where the subscript of each e indicates a square in the grid.

   Thus the agent, sitting in the bottom lefthand corner, is in state $e_{0,0}$, while if the agent were at the goal, in the top righthand corner, it would be in state $e_{5,5}$. $e_{0,0}$ is the initial state of the environment.

   The filled squares indicate obstacles — the agent cannot be in these states.

   The agent can move north, south, east or west, which we write as:

   $$\alpha_n, \alpha_s, \alpha_e, \alpha_w$$

   and these have the effects you would expect. If the agent is in state $e_{0,0}$ and takes action α_n, it will end up in state $e_{0,1}$, while if the agent is in state $e_{3,2}$ and takes action α_e it will end up in $e_{4,2}$. If the agent tries to move outside the grid then it does not move (for example if the agent is in $e_{0,5}$ and tries to do α_n then it stays in $e_{0,5}$).

   If the agent enters the state $e_{5,5}$, marked with the word goal, then it gets a reward of 10. If it enters state $e_{1,4}$, marked with a dark circle, it gets a reward of -10 (ie it takes a loss).

   (a) Give one *maintenance* task for the agent in this environment and one *achievement* task.
   (10 points)

   (b) Write down a run of the agent that takes it from $e_{0,0}$ to $e_{1,4}$.
   (10 points)

   (c) Consider the following control program:

   while(not in state $e_{5,5}$){
       randomly pick either α_n or α_e (each with probability 0.5)
       execute the action that was selected
   }

   Write down two runs that the agent might carry out when executing this program.
   (10 points)

   (d) If the agent executes the program, can it ever reach $e_{5,0}$? Why?
   (10 points)

   (e) What is the maximum and minimum reward that the agent can get running this program? Why?
   (10 points)

   (f) How likely is it that when the agent runs the above program it will get the reward of -10?
   (Hint: There is a precise probability that the agent will get to the relevant state. That is what I want you to calculate).
   (10 points)

2. Suppose the vacuum world from Lecture #3 contains obstacles which the agent has to avoid, and that the agent has a sensor to detect the obstacles. Write down a new logic-based solution for the agent.

   Explain what, if any, guarantees you can make for your solution about its ability to make the room clean, and justify your answer.

   (20 points)

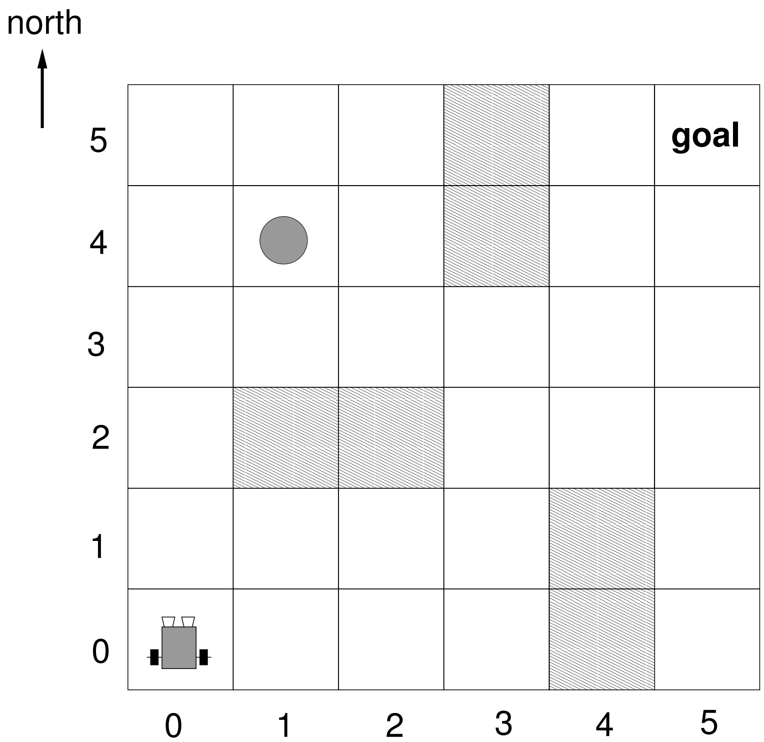

Figure 1:

3. Suppose that the vacuum world agent's sensors — the one from the lecture that sees the dust, and the new one from Question 2 that can see obstacles — are now *noisy* so they only give the right answer 80% of the time.

   (You can interpret this to mean that if there is dust or an obstacle, the relevant sensor will say that it is there 80% of the time. You can assume when there is no dust or obstacle, the sensor always correctly reports this.)

   How does this change the logic-based control program? Assume that if the agent tries to move into a square that contains an obstacle, it does not move.

   Explain what, if any, guarantees you can make for your solution about its ability to make the room clean, and justify your answer.

   (20 points)