COMP-424: Artificial intelligence

Homework 4

Due on *myCourses* Thursday Apr 6, 11:59pm.

General instructions.

- This is an individual assignment. You can discuss solutions with your classmates, but should only exchange information orally, or else if in writing through the discussion board on *myCourses*. All other forms of written exchange are prohibited.
- Unless otherwise mentioned, the only sources you should need to answer these questions are your course notes, the textbook, and the links provided. Any other source used should be acknowledged with proper referencing style in your submitted solution.
- Submit a single pdf document containing all your pages of your written solution on your McGill's *myCourses* account. You can scan-in hand-written pages. If necessary, learn how to combine many pdf files into one.

Question 1: Hidden Markov Models

Consider a HMM used by a hospital to track the well-being of a patient. We will consider a simple case in which the patient is described as being in one of two states: {*Well*, *Unwell*}, and there are two observations corresponding to the patient's blood pressure: {*Normal*, *Abnormal*}.

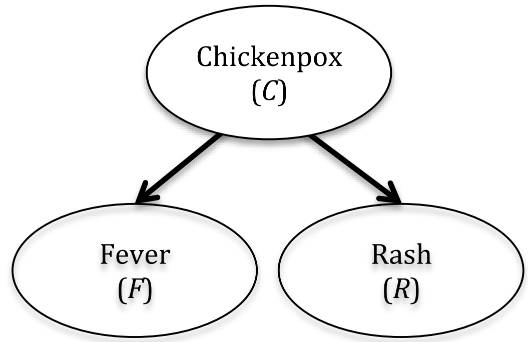
The prior probability of the patient being unwell, Pr(*Unwell*) = 0.1. At each time step, if the patient is unwell, there is a 0.5 probability that the patient will continue to be unwell in the next time step. If the patient is well, there is a 0.8 probability that the patient will stay well in the next time step. The probability of observing a normal blood pressure if the patient is well is 0.6. The corresponding probability if the patient is unwell is 0.3.

a) Draw the HMM that describes this scenario, clearly indicating the relevant parameters using conditional probability tables.
b) What is the probability of the observation sequence {*Abnormal*, *Normal*, *Normal*}?
c) What is the probability of the patient being unwell after observing the sequence in b)?
d) What is the most likely sequence of three states explain the observation sequence in b)?

For each of these tasks, write down your calculations, or provide the code that you wrote to compute the answer.

Question 2: Utility

Consider the Bayes Net shown here, with all Bernoulli variables, which models the chickenpox disease. Assume $P(C)=0.05$, $P(F|C)=0.6$, $P(F|\neg C)=0.1$, $P(R|C)=0.9$, $P(R|\neg C)=0.2$. Add the fact that having chickenpox has a utility of -100 if detected and properly treated with anti-viral medication, but has a utility of -200 if left untreated (the Utility doesn't change as a function of Fever or Rash). Applying treatment when there is no virus has a utility of -20, and applying no treatment if there is no virus has a utility of 0.

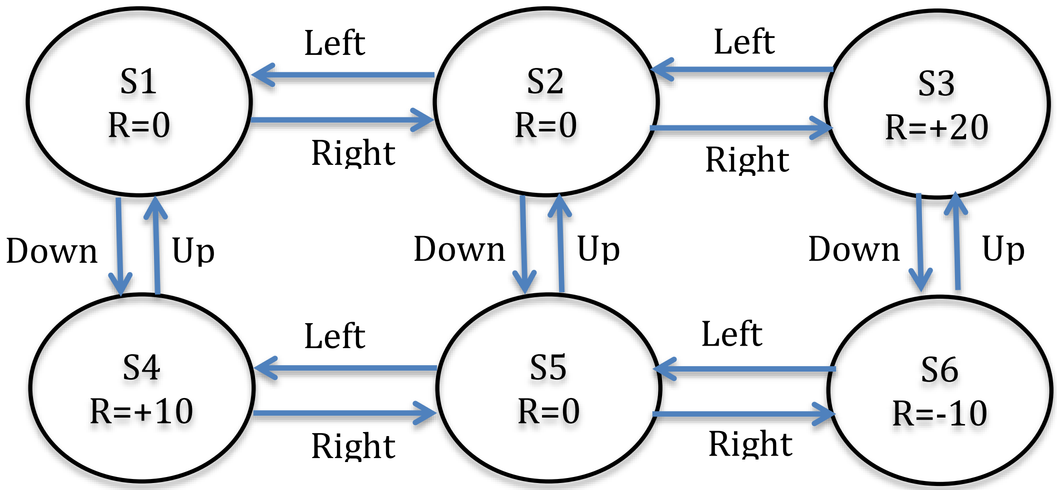


Use the principle of Maximum Expected Utility and Value of Information to answer the following questions. For parts a)-c) assume the chickenpox diagnostic test has 100% accuracy.

a) Given no information on fever or rash, how much should you pay to find out whether or not you have chickenpox?
b) How much should you pay to find out whether or not you have chickenpox if you observe a fever (*F=True*) and but no rash (*R=False*)?
c) How much should you pay to find out whether or not you have chickenpox if you observe a rash (*R=True*), but don't know whether or not you have a fever?
d) A company has developed a new test for chickenpox, however this test is not perfectly accurate: If the chickenpox virus is present, it correctly finds it with 80% accuracy but fails to find it in 20% of tests, and if the chickenpox virus is not present, it reports a false positive in 20% of cases. How much should they charge for this test, to make it competitive with the diagnostic test that has 100% accuracy? (*Hint*: Set the cost of the new test to have the same MEU as the other test.)

Question 3: Markov Decision Processes

Consider the MDP shown below. It has 6 states and 4 actions. As shown on the figure, the transitions for all actions have a Pr=0.7 of succeeding (and leading to the state shown by the arrow) and Pr=0.3 of failing (in which case the agent stays in place). For other transitions that are not shown, assume that they cause the state to stay the same (e.g. $T(S1,Left,S1)=1$). The rewards depend on state only and are shown in each node (state); rewards are the same for all actions (e.g. $R(S4,a)=+10,\ \forall a$). Assume a discount factor of $\gamma=0.9$.



a) Describe the space of all possible policies for this MDP. How many are there?
b) Assuming an initial policy $\pi^0(s)=Right,\ \forall s$, perform policy evaluation to get the initial value function for each state, $V^0(s)$, $\forall s$.
c) Given the initial estimate, V^0, if you run an iteration of policy improvement, what will be the new policy at each state? If necessary, break ties alphabetically, e.g. "Down" before "Left", etc.)
d) What is the optimal value function at each state for this domain?
e) Is the optimal value function unique? Explain.
f) What is the optimal policy at each state for this domain?
g) Is the optimal policy unique? Explain.
h) Suggest a change to the reward function that changes the value function but does not change the optimal policy.

Question 4: Reinforcement Learning

(Reproduced from Russell & Norvig, Exercise 21.4, pg. 858)

Write out the parameter update equations for TD learning with

$$\hat{U}(x,y) = \theta_0 + \theta_1 x + \theta_2 y + \theta_3 \sqrt{(x - x_g)^2 + (y - y_g)^2}.$$