CSE 6369 - *Reasoning with Uncertainty*

Homework 2- Fall 2015

Due Date: Nov. 12 2015, 11:59 pm

Monte-Carlo Estimation

1. You are given the following Markov process and observation model:
   State space: $S = \{1, 2, 3, 4, 5, 6\}$

   Transition probabilities: $P(s|s') = T_{s,s'}, T = \begin{pmatrix} 0.7 & 0.15 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.3 & 0.7 & 0.15 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.15 & 0.7 & 0.15 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.15 & 0.7 & 0.15 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.15 & 0.7 & 0.3 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.15 & 0.7 \end{pmatrix}$

   Observations: $O = \{a, b, c, d\}$

   Observation probabilities: $P(o|s) = B_{o,s}, B = \begin{pmatrix} 0.7 & 0.1 & 0.0 & 0.0 & 0.1 & 0.7 \\ 0.2 & 0.7 & 0.1 & 0.1 & 0.7 & 0.2 \\ 0.1 & 0.2 & 0.4 & 0.4 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.5 & 0.5 & 0.0 & 0.0 \end{pmatrix}$

   a) Show three iterations of a particle filter with 15 samples for the following conditions:
   Start state: 4 (with probability 1)
   Observed sequence (the first observation is made after the first time step - i.e. after the first model update): d, c, b, a

   b) Use the result of your filter to determine an estimate for the expected value of the properties $f(s)$ and $g(s)$ (interpreting the numbers representing states as continuous values) after 4 model steps. $f(s) = s$ (the mean of the distribution), $g(s) = (s - \hat{s})^2$, where $\hat{s}$ is the expected value for the state in the given distribution (i.e. $E[g(s)]$ is (approximately) the variance of the distribution).

Hidden Markov Models

2. You are given the following Hidden Markov Models for two differently biased coins:

Model 1:

State space: $S_1 = \{s_1, s_2\}$
Observations: $O_1 = \{H, T\}$
Transition probabilities: $P_1(s_i|s_j) = T_{i,j}, T = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$
Observation probabilities: $P_1(o|s) = B_{o,s}, B = \begin{pmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{pmatrix}$
Prior probabilities: $\pi_1(s) = \Pi_s, \Pi = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$

Model 2:

State space: $S_2 = \{s_1, s_2\}$
Observations: $O_2 = \{H, T\}$
Transition probabilities: $P_2(s_i|s_j) = T_{i,j}, T = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix}$
Observation probabilities: $P_2(o|s) = B_{o,s}, B = \begin{pmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{pmatrix}$
Prior probabilities: $\pi_2(s) = \Pi_s, \Pi = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}$

Use the Forward algorithm to determine which of the two models better describes an actual coin used which produced the observation sequence H,T,H,H,H,H,H,T,H,T. List all the α values involved in the calculations[1].

3. Build a Hidden Markov Model for the following (unrealistically simplified) intrusion detection problem: The state of a given network node is either uncompromised or hijacked and the only observable information about the node is the average number of accesses to a particular port. This information is provided once per hour and based on past experience, it falls into 4 groups: a : 0?10, b : 10?20, c : 20?30, and d : 30?40. It was also observed that for an uncompromised node these observations are made with the following relative frequencies: a : 60% of the time, b : 20% of the time, c : 10% of the time, d : 10% of the time. For a hijacked node the observations were made with the following relative frequencies: a : 20% of the time, b : 10% of the time, c : 40% of the time, d : 30% of the time. Further study of past experience revealed that a node is being hijacked on average every 8 hours and is used by the attacker for an average of 3 hours before being released.

   a) Build a Hidden Markov Model that represents the scenario described above. Assume that at the beginning of the observations the node is uncompromised.

[1] Rather than computing all the values by hand it might be simpler to write a short program to compute the values.

b) Given the following observation sequence, $a, b, c, a, b, d, c, d, c, d, a, b$, taken between $6:00$ pm and $5:00$ am, determine if the node has likely been hijacked by determining the most likely interpretation (i.e. corresponding state sequence) using the Viterbi algorithm. List the intermediate values (δ and Ψ) of the Viterbi algorithm[1].