Homework Set #6

Due: 11/12/03

1 Problem: Bayes ball

In the class we discussed the notion of how local Bayesian network properties impact global distribution described by the network. One way to assess dependencies/independencies between nodes in the network is through so called "Bayes ball" rules [Shachter98]. These rules rely on atomic network structures depicted in Figure 1.

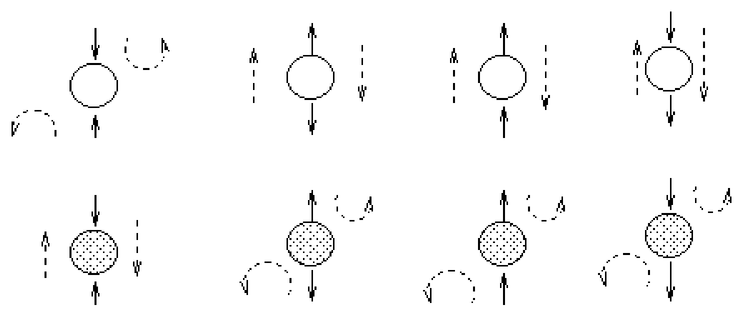

Figure 1: Bayes ball rules

The rules are as follows:

Two (sets of) nodes A and B are conditionally independent (d-separated) given a set of nodes C if and only if there is no way for a ball to get from A to B in the graph, where the allowable movements of the ball are shown in Figure 1. Hidden nodes are nodes whose values are not known, and are depicted as unshaded; observed nodes (the ones we condition on) are shaded. The dotted arcs indicate direction of flow of the ball.

1. Column 1: **Converging arcs**. If two arcs are converging on a node (call it C), and C **IS NOT** instantiated then the ball bounces off of node C. On the other hand, if C **IS** instantiated, the ball can pass through.

2. Column 2: **Diverging arcs**. If two arcs are diverging from a node (C), then the ball will bounce off of node C when C **IS** instantiated. If C **IS NOT** instantiated, the ball passes through.

3. Columns 3 & 4: **Incoming-Outgoing arcs**. If one of the arcs is incoming and the other one is outgoing, the ball will bounce off if node C **IS** instantiated. If C **IS NOT** instantiated, the ball passes through.

Consider the case when each of the two arcs in these four atomic structures has a node attached to it and call them A (top) and B (bottom). Prove the three Bayes ball rules using the rules of probability (Bayesian rule and others). For instance, for Converging arcs rule you need to prove that $P(A,B) = P(A)P(B)$ when C is not instantiated and $P(A,B|C) \neq P(A|C)P(B|C)$ when C is instantiated.

2 Problem: Bayesian network and Hugin

In this problem you will need to use Hugin Expert Lite (http://www.hugin.com/Products_Services/Products/Demo/Lite). Go to the web site and download one of the distributions available (Windows, Sun, or Linux). You will also need a copy of the Bayesian network known as the Asia net, which is available at http://developer.hugin.com/Samples/Asia/Asia.article.

Asia is a small Bayesian network that calculates the probability of a patient having tuberculosis, lung cancer or bronchitis respectively based on different factors - for example whether or not the patient has been to Asia recently. Shortness-of-breath (dyspnoea) may be due to tuberculosis, lung cancer, bronchitis, more than one of these diseases or none of them. A recent visit to Asia increases the risk of tuberculosis, while smoking is known to be a risk factor for both lung cancer and bronchitis. The results of a single chest X-ray do not discriminate between lung cancer and tuberculosis, as neither does the presence or absence of dyspnoea.

1. For each node in the network list the nodes in its Markov blanket.

2. Using the Bayes ball rules comment on dependency/independency among $Has\_tuberculosis$, Has_lung_cancer, and $Has\_bronchitis$ in the following situations:

   - $Dyspnoea$ is known.
   - $Dyspnoea$ and $PositiveXray$ are known.
   - $Dyspnoea$, $PositiveXray$, and $VisitToAsia$ are known.
   - $Dyspnoea$, $PositiveXray$, $VisitToAsia$, and $Smoker$ are known.

3. Simulate 1000 samples of this network using Hugin. (You do not need to submit a record of those 1000 samples—they will be used in (4).)

4. From the simulated set, compute the following probabilities, using (a) rejection sampling and (b) likelihood weighting:

   - $Pr(Has\_lung\_cancer)$
   - $Pr(Has\_tuberculosis|Dyspnoea = True)$.
   - $Pr(Has\_tuberculosis|Dyspnoea = True, PositiveXray = True)$.
   - $Pr(Has\_tuberculosis|Dyspnoea = True, PositiveXray = True, VisitToAsia = False)$.
   - $Pr(Has\_tuberculosis|Dyspnoea = True, PositiveXray = True, VisitToAsia = False, Smoker = True)$.

5. Now use Hugin to compute the same probabilities using "Propagate Sum-Normal" inference mode. Compare the two sets of probabilities and explain why or why not there are any discrepancies.

6. Assume that the physician who evaluates this patient failed most of his medical school exams and cannot accurately observe patient's symptoms. In fact, given a true symptom such as $PositiveXray$, the probability that the physician made the same diagnosis is 50%. Introduce the two new nodes in the network, called $EvaluatedPositiveXray$ and $EvaluatedDyspnoea$ that correspond to these physician evaluations. For instance, $EvaluatedPositiveXray$ will be connected to $PositiveXray$, with an arc directed toward $EvaluatedPositiveXray$. Now compute the following probabilities (using Hugin):

   - $Pr(HasLungCancer)$
   - $Pr(Tuberculosis|EvaluatedDyspnoea = True)$.
   - $Pr(Tuberculosis|EvaluatedDyspnoea = True, EvaluatedPositiveXray = True)$.
   - $Pr(Tuberculosis|EvaluatedDyspnoea = True, EvaluatedPositiveXray = True, VisitToAsia = False)$.
   - $Pr(Tuberculosis|EvaluatedDyspnoea = True, EvaluatedPositiveXray = True, VisitToAsia = False, Smoker = True)$.

   Do they differ from the probabilities computed in (5) and, if so, why?

Table 1: HMM parameters.

Initial state probability vector	$\pi_0 = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$
State transition probability matrix	$\Pi = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}$
Emission probability matrix	$Q = \begin{bmatrix} .7 & .2 \\ .1 & .25 \\ .2 & .55 \end{bmatrix}$

3 Problem: HMM

You are given a two-state, $S_k \in \{1, 2\}$, discrete HMM with three observation symbols, $Y_k \in \{A, B, C\}$, with the parameters shown in Table 1. In other words, e.g., $Pr(S_1 = 1) = .5$, $Pr(S_k = 1|S_{k-1} = 2) = .5$ and $Pr(Y_k = A|S_k = 1) = .7$.

Assume you observed the following sequence of symbols: $Y_{1:7} = \{A, A, B, C, C, B, C\}$. Answer the following questions:

1. What is the probability of this sequence, given the above model, i.e., $Pr(Y_{1:7}|model)$.
2. Find the most likely sequence of states, $S^*_{1:7} = \arg\max_{S_{1:7}} Pr(S_{1:7}|Y_{1:7}, model)$.
3. Find the posterior distribution of states at position $k = 4$ given the sequence $Y_{1:7}$, $Pr(S_4|Y_{1:7}, model)$.
4. Find the posterior distribution of states at position $k = 4$ given a partial sequence $Y_{1:4} = \{A, A, B, C\}$, $Pr(S_4|Y_{1:4}, model)$.

4 Problem: Kalman filters

Consider the problem of modeling a simple Newtonian motion of a point (object) mass described by the Newton law of dynamics

$$m \frac{dv}{dt} = F$$

where m is the mass of the object, v is its velocity, and F is the force excerpted on the object. Let us call this Equation 1.

1. Let $x(t)$ be the position of the object at time t. How is the object's position related to its velocity? Call this Equation 2.
2. Approximate continuous derivatives in Equations 1 & 2 using simple finite differences, e.g., $dv/dt = (v_t - v_{t-\Delta})/\Delta$. Now write the object motion equations in the following form:

$$\begin{bmatrix} x_t \\ v_t \end{bmatrix} = A \begin{bmatrix} x_{t-\Delta} \\ v_{t-\Delta} \end{bmatrix} + b,$$

   where A is a matrix and b is a vector. Show exact expressions for A and b.
3. Consider the case of $F = 0$. Let $v_0 = 2$, $x_0 = 0$, and $\Delta = 1$. Using the equation from (2) compute the values of x_t and v_t for $t = 0, 1, \ldots, 10$.
4. Assume that we add some Gaussian noise to both x_t and v_t with mean 0 and variance .1, for all $t = 0, \ldots, 10$. Generate a sequence of such noisy versions of x_t and v_t. Call them z_t and y_t (i.e., $z_t = x_t + noise$.)
5. Write the equations for x_t, v_t and z_t in the form suitable for Kalman filtering, i.e.,

$$\begin{bmatrix} x_t \\ v_t \end{bmatrix} = A \begin{bmatrix} x_{t-\Delta} \\ v_{t-\Delta} \end{bmatrix}$$

$$z_t = C \begin{bmatrix} x_t \\ v_t \end{bmatrix} + noise$$

6. Now assume that you are only given the noisy positional measurements of the object, z_t, from (4). Using Kalman filtering equations, compute the estimates of the mean and variance of $x_t | Z_{0:t}$.