Complete and hand in the following exercises before 31st of January, 2003 to the course assistant (Room C311 in the computer science building). The exercise will affect your grade on the course. The solutions will be presented at a later date, most probably during February, 2003.

1. Let's assume that there are three types of possible weather: *cloudy, sunny*, and *rainy*. On each day, the same weather lasts always from morning to night. Let's assume that the weather develops according to a first-order Markov chain. The weather transition matrix is as follows:

| | | Tomorrow's weather | | |
|---|---|---|---|---|
| | | Cloudy | Sunny | Rainy |
| Today | Cloudy | 0.5 | 0.2 | 0.3 |
| | Sunny | 0.25 | 0.7 | 0.05 |
| | Rainy | 0.4 | 0.2 | 0.4 |

   What we want to do is to try to infer the future weather based on the observed weather history. For some reason, you're sitting inside and cannot actually observe the weather (the weather is hidden from you). You do, however, observe people with or without rain gear each day and use this information for weather prediction.

   Then, suppose that the probabilities for seeing people with rain gear on a given weather are:

| Weather | Prob. of rain gear |
|---|---|
| Cloudy | 0.3 |
| Rainy | 0.85 |
| Sunny | 0.05 |

   The day you actually last saw the weather was a rainy day. The next day (day 2) you did not see people with rain gear but on day 3 you did see. What is the probability that it is sunny on the third day?

2. Draw the graphical representation of a Bayesian network in which the following conditional independencies hold:

$$P(y_0, \ldots, y_T, s_0, \ldots, s_T) = P(s_0) \prod_{t=0}^{T} P(y_t|s_t) \prod_{t=1}^{T} P(s_t|s_{t-1})$$

   Derive the junction tree for the model. What operations are required for the derivation of the junction tree for this particular model? Derive the iterative update rules for inferring $P(s_t|y_1, \ldots, y_t)$, $t = 1, \ldots, T$.

3. A mixture model is shown in Figure 1. Give the joint distribution of the variables $P(S,Y)$ and the probability of observed data given the variable $S$ as in $P(Y|S)$. How many variables does the largest clique of the junction tree derived for this model have? Are $y_1$ and $y_2$ married when deriving the moral graph? What can be said about $S$ and the separators in the junction tree?

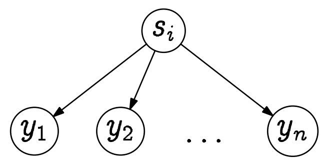


Figure 1: A mixture model.

4. Look at the junction tree of Figure 2. Name the parts of the junction tree and give the condition that affirms the running intersection property for this particular junction tree. Which Bayesian network(s) may be represented with this tree. Draw the graphical representation of the models.

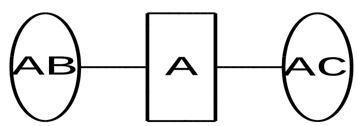


Figure 2: A simple junction tree.

5. Assume a simple Bayesian network in Figure 3. First, calculate $P(A|C)$ using the Bayes's rule. Next, form the junction tree and insert evidence $A = a_1$. List the local messages and the probability distribution at the globally consistent state.

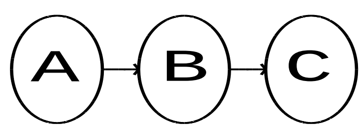


Figure 3: A Bayesian network.

6. Considering the more complicated graph structure with two hidden layers, as shown in Figure 4.

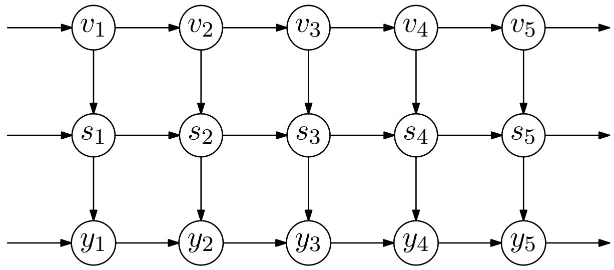


Figure 4: A three layer Markov model.

Assuming that the variables of the model $P(V, S, Y)$ can only take binary values ($y_t, s_t, v_t \in \{0, 1\}$), give the probabilities for each of the variables at time $t$ conditioned on the previous states of the variables.

Give the inference algorithms that can be used for estimating the smoothed probabilities of the hidden states $P(v_t, s_t|Y_T)$ after all data $Y = \{y_1, y_2, \ldots, y_T\}$ have been observed. Start with the filtering part, that is, estimation of $P(v_t, s_t|Y_t)$ when data $Y = \{y_1, \ldots y_t, \ldots, y_T\}$ is available.