CS340 Fall 2006: Homework 8

Out Fri 17 Nov, back Fri 24 Nov

1 Bayes Ball

Here we compute some global independence statements from some directed graphical models using the "Bayes ball" algorithm (equivalent to d-separation).

1. Consider the DAG in Figure 1(a). Determine which variables are d-separated from A given that B is observed (i.e., list all variables X s.t., $X \not\perp A|B$). (Shaded nodes, with lines through them, are observed.)

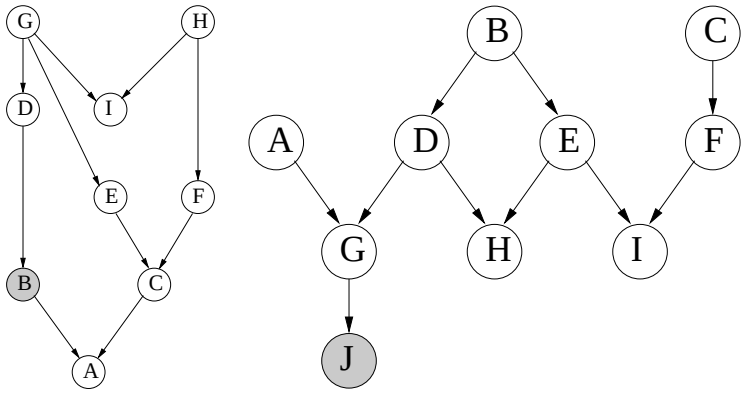
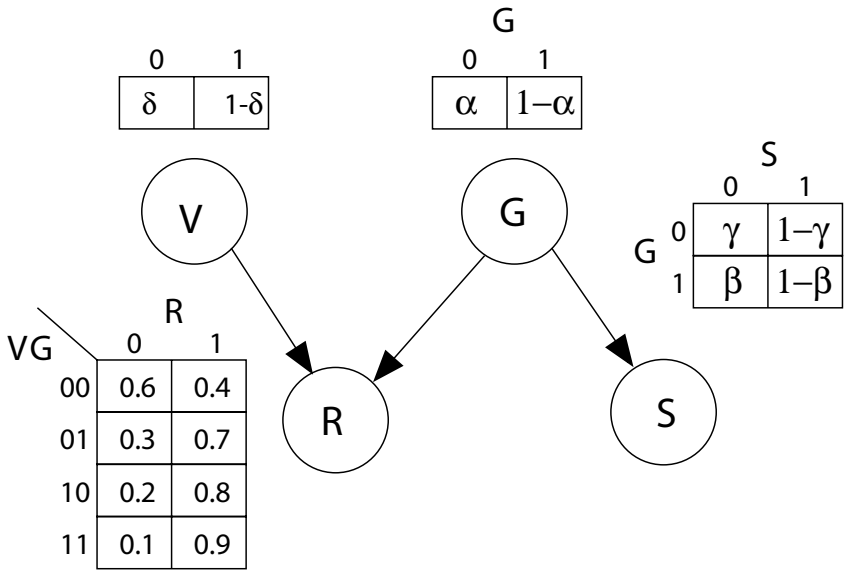


Figure 1: Bayes nets

2. Consider the DAG in Figure 1(b). Determine which variables are d-separated from A given that J is observed (i.e., list all variables X s.t., $X \not\perp A|J$). (Shaded nodes, with lines through them, are observed.)

2 Bayes nets for a rainy day

In this question you must model a problem with 4 binary variables: $G =$"gray", $V =$"Vancouver", $R =$"rain" and $S =$"sad". You are given the following graphical model describing the relationship between these variables:



1. Write down an expression for $P(S = 1|V = 1)$ in terms of $\alpha, \beta, \gamma, \delta$.
2. Write down an expression for $P(S = 1|V = 0)$. Is this the same or different to $P(S = 1|V = 1)$? Explain why.
3. Find maximum likelihood estimates of α, β, γ using the following data set, where each row is a training case. (You may state your answers without proof.)

$$\begin{array}{cccc} V & G & R & S \\ \hline 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{array} \tag{1}$$

4. What is wrong with these maximum likelihood estimates? What is an easy way to fix this problem?

3 Fishing nets

Consider the Bayes net shown in Figure 2. Here, the nodes represent the following variables

$$\begin{aligned} X_1 &\in \{\text{winter, spring, summer, autumn}\} \\ X_2 &\in \{\text{salmon, sea bass}\} \\ X_3 &\in \{\text{light, medium, dark}\} \\ X_4 &\in \{\text{wide, thin}\} \end{aligned}$$

The corresponding conditional probability tables are

$$p(x_1) = \begin{pmatrix} .25 & .25 & .25 & .25 \end{pmatrix}$$

$$p(x_2|x_1) = \begin{pmatrix} .9 & .1 \\ .3 & .7 \\ .4 & .6 \\ .8 & .2 \end{pmatrix}$$

$$p(x_3|x_2) = \begin{pmatrix} .33 & .33 & .34 \\ .8 & .1 & .1 \end{pmatrix}$$

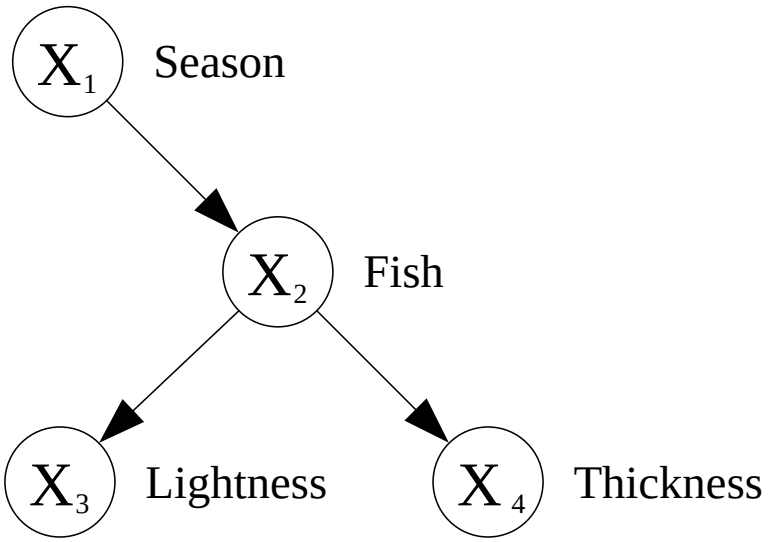


Figure 2: Fish Bayes net

$$p(x_4|x_2) = \begin{pmatrix} .4 & .6 \\ .95 & .05 \end{pmatrix}$$

Note that in $p(x_4|x_2)$, the rows represent x_2 and the columns x_4. Thus $p(x_4 = \text{thin}|x_2 = \text{sea bass}) = 0.05$, etc. Answer the following queries. You may use matlab or do it by hand. In either case, show your work.

1. Suppose the fish was caught on December 20 — the end of autumn and the beginning of winter — and thus let $p(x_1) = (.5,\ 0,\ 0,\ .5)$ instead of the above prior. (This is called **soft evidence**, since we do not know the exact value of X_1, but we have a distribution over it.) Suppose the lightness has not been measured but it is known that the fish is thin. Classify the fish as salmon or sea bass.

2. Suppose all we know is that the fish is thin and medium lightness. What season is it now, most likely? Use $p(x_1) = \begin{pmatrix} .25 & .25 & .25 & .25 \end{pmatrix}$