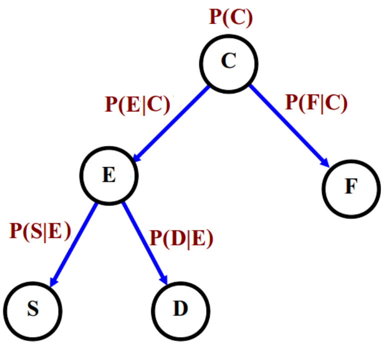
Tutorial 1: Bayesian Decision Trees, λ and π evidence

In lecture 2 we introduced a naive Bayesian network for recognising cats in pictures:



Variable	Interpretation	Value
C	Cat	c_1, c_2
E	Eyes	e_1, e_2, e_3
S	Separation of the eyes	$s_1, s_2, s_3, s_4, s_5, s_6, s_7$
D	Difference in eye size	d_1, d_2, d_3, d_4
F	Fur colour	$f_1, f_2, \cdots f_{10}$

Suppose that for a given data set we obtain the following link matrices:

$$P(D|E) = \begin{bmatrix} P(d_1|e_1) & P(d_1|e_2) & P(d_1|e_3) \\ P(d_2|e_1) & P(d_2|e_2) & P(d_2|e_3) \\ P(d_3|e_1) & P(d_3|e_2) & P(d_3|e_3) \\ P(d_4|e_1) & P(d_4|e_2) & P(d_4|e_3) \end{bmatrix} = \begin{bmatrix} 0.4 & 0.33 & 0.29 \\ 0.4 & 0.33 & 0.14 \\ 0.2 & 0.34 & 0.14 \\ 0 & 0 & 0.43 \end{bmatrix}$$

$$P(S|E) = \begin{bmatrix} P(s_1|e_1) & P(s_1|e_2) & P(s_1|e_3) \\ P(s_2|e_1) & P(s_2|e_2) & P(s_2|e_3) \\ P(s_3|e_1) & P(s_3|e_2) & P(s_3|e_3) \\ P(s_4|e_1) & P(s_4|e_2) & P(s_4|e_3) \\ P(s_5|e_1) & P(s_5|e_2) & P(s_5|e_3) \\ P(s_6|e_1) & P(s_6|e_2) & P(s_6|e_3) \\ P(s_7|e_1) & P(s_7|e_2) & P(s_7|e_3) \end{bmatrix} = \begin{bmatrix} 0 & 0.33 & 0.14 \\ 0.6 & 0 & 0.14 \\ 0.4 & 0.34 & 0 \\ 0 & 0.33 & 0.14 \\ 0 & 0 & 0.14 \\ 0 & 0 & 0.14 \\ 0 & 0 & 0.28 \end{bmatrix}$$

$$P(F|C) = \begin{bmatrix} P(f_1|c_1) & P(f_1|c_2) \\ P(f_2|c_1) & P(f_2|c_2) \\ P(f_3|c_1) & P(f_3|c_2) \\ P(f_4|c_1) & P(f_4|c_2) \\ P(f_5|c_1) & P(f_5|c_2) \\ P(f_6|c_1) & P(f_6|c_2) \\ P(f_7|c_1) & P(f_7|c_2) \\ P(f_8|c_1) & P(f_8|c_2) \\ P(f_9|c_1) & P(f_9|c_2) \\ P(f_{10}|c_1) & P(f_{10}|c_2) \end{bmatrix} = \begin{bmatrix} 0 & 0.3 \\ 0.125 & 0 \\ 0.125 & 0.14 \\ 0.25 & 0.14 \\ 0.125 & 0 \\ 0.125 & 0 \\ 0 & 0.14 \\ 0.125 & 0.14 \\ 0 & 0.14 \\ 0.125 & 0 \end{bmatrix}$$

$$P(E|C) = \begin{bmatrix} P(e_1|c_1) & P(e_1|c_2) \\ P(e_2|c_1) & P(e_2|c_2) \\ P(e_3|c_1) & P(e_3|c_2) \end{bmatrix} = \begin{bmatrix} 05 & 0.14 \\ 0.25 & 0.14 \\ 0.25 & 0.72 \end{bmatrix}$$

The prior probability of finding a cat in an image is taken as $P(C) = [0.5, 0.5]$.

P.T.O.

1. An image is processed, and the following data extracted: $[s_1, d_2, f_3]$. Calculate the λ evidence that is propagated and the probability distributions over C.

2. The same image is processed in monochrome. Thus there is no information for F. Re-calculate the probabilities of C for the cases given in question 1. Remember that the data is $[s_1, d_2]$, and since nothing is known about F, the λ value for every state of F is set to 1.

3. Doubt is expressed about the accuracy of the computer vision algorithms. Thus instead S and D are instantiated with virtual evidence as follows:

State	$\lambda(s_i)$
s_1	0.8
s_2	0.2
s_3	0.0
s_4	0.0
s_5	0.0
s_6	0.0
s_7	0.0

State	$\lambda(d_i)$
d_1	0.3
d_2	0.4
d_3	0.3
d_4	0.0

   There is still no information for node F.

   Re-calculate the distributions over C.

4. Given the evidence for C in part 3, calculate the probability distribution over F.