Solutions for Bayesian networks and decision graphs (second edition)

Finn V. Jensen and Thomas D. Nielsen

August 23, 2007

Solution for exercise 3.1

There are two information variables, *English Grade* and *Math Grade* and three hypothesis variables, *Prob. Grade*, *Ling. Grade*, and *Alg. Grade*. Mediating variables could be *Mathematical Talent* and *Linguistic Talent*.

Solution for exercise 3.2

See the Hugin network, where a suggestion for the structure is given. An extra variable is introduced for modelling your ability to actually spot the spots.

Solution for exercise 3.3

See the Hugin network.

Solution for exercise 3.4

Solution for exercise 3.5

See the Hugin network.

Solution for exercise 3.6

See the Hugin network.

Solution for exercise 3.7

Solution for exercise 3.8

Part 1:

$$P(Pr|Bt=n,Ut=n)=\frac{P(Pr,Bt=n,Ut=n)}{P(Bt=n,Ut=n)}$$

Next, we should calculate $P(Pr, Bt = n, Ut = n)$ and $P(Bt = n, Ut = n)$:

$$P(Pr,Bt=n,Ut=n)=\sum_{Ho}P(Pr,Bt=n,Ut=n,Ho)$$
$$P(Bt=n,Ut=n)=\sum_{Pr}P(Pr,Bt=n,Ut=n)$$

Thus, we only need to calculate $P(Pr, Bt = n, Ut = n, Ho)$ and this can be done using the chain rule:

$$P(Pr,Bt=n,Ut=n,Ho)=P(Ut=n|Ho)P(Bt=n|Ho)P(Pr|Ho)P(Pr)$$

The final value is $P(Pr|Bt = n, Ut = n) = (0.53, 0.47)$, see also the Hugin network.

Part 2:

The following probabilities can be calculated from the original network by inserting the evidence $Pr = y$ and $Pr = n$, respectively:

$$P(BT|Pr=y)=(0.64,0.36)$$
$$P(BT|Pr=n)=(0.106,0.894)$$
$$P(UT|Pr=y)=(0.73,0.27)$$
$$P(UT|Pr=n)=(0.107,0.893)$$

$P(Pr|Bt = n, Ut = n) = (0.449, 0.551)$, see also the Hugin network. Notice the difference between this result and the result you got from the original network.

Solution for exercise 3.9

See the Hugin network.

Solution for exercise 3.10

(i) See the Hugin network.

(ii) Out of 10,000 we every day get two infected, and as the number shall be stable, two of the seven must be cured. This is not quite exact. The exact number is $0.9993 \times \frac{2}{7}$.

(iii) The same (approximate reasoning as above).

Solution for exercise 3.11

According to the chain rule for Bayesian networks we have

$$P(H, I_1, ..., I_n) = P(H) \prod_i P(I_i|H)$$

then

$$P(H, f_1, ..., f_n) = P(H) \prod_i P(f_i|H)$$

and

$$P(H|f_1, ..., f_n) = \frac{P(H)}{P(f_1, ..., f_n)} \prod_i P(f_i|H).$$

Solution for exercise 3.12

See the Hugin network.

Solution for exercise 3.13

See the Hugin network.

Solution for exercise 3.14

See the Hugin network.

Solution for exercise 3.15

See the Hugin network. For the model in Exercise 3.14, the evidence $FC = 1, SC = 2$ in inconsistent. The naive Bayes model yields

$$P(OH2|FC = 1, SC = 2) = (36.75, 31.84, 16.81, 5.71, 5.7, 2.59, 0.32, 0.27).$$

Solution for exercise 3.16

See the Hugin network.

Solution for exercise 3.17

See the Hugin network.

Solution for exercise 3.18

See the Hugin network.

Solution for exercise 3.19

See the Hugin network.

Solution for exercise 3.20

For notational convenience we assume that the causes which are present are indexed from 1 to i, and that the remaining causes are off. That is, we have:

$$P(B = y|A_1 = \cdots A_i = y, A_{i+1} = \cdots = A_n = n) = \\ 1 - P(B = n|A_1 = \cdots A_i = y, A_{i+1} = \cdots = A_n = n).$$

From the chain rule (and marginalization) we get:

$$P(B = n|A_1 = \cdots A_i = y, A_{i+1} = \cdots = A_n = n) = \\ \sum_{B_1,\ldots,B_n} P(B = n|B_1, \ldots, B_n) \prod_{j=1}^{i} P(B_j|A_j = y) \prod_{k=i+1}^{n} P(B_k|A_k = n).$$

From the construction of $P(B = n|B_1, \ldots, B_n)$ (which specifies a logical-or) we have that $P(B = n|B_1, \ldots, B_n) = 0$ if there is a B_i s.t. $B_i = y$. That is, the only configuration which will contribute with a value different from 0 is $(B_1 = n, \ldots, B_n = n)$. We can therefore write:

$$P(B = n|A_1 = \cdots A_i = y, A_{i+1} = \cdots = A_n = n) = \\ P(B = n|B_1 = n, \ldots, B_n = n) \prod_{j=1}^{i} P(B_j = n|A_j = y) \\ \prod_{k=i+1}^{n} P(B_k = n|A_k = n),$$

and since $P(B = n|B_1 = n, \ldots, B_n = n) = 1$ and $P(B_k = n|A_k = n) = 1$, for all $k \geq i+1$ we have:

$$P(B = n | A_1 = \cdots A_i = y, A_{i+1} = \cdots = A_n = n)$$
$$= \prod_{j=1}^{i} P(B_j = n | A_j = y) = \prod_{j=1}^{i} q_j$$

and therefore

$$P(B = y | A_1 = \cdots A_i = y, A_{i+1} = \cdots = A_n = n) = 1 - \prod_{j=1}^{i} q_j$$

which completes the proof.

Solution for exercise 3.21

Part 1: The proof follows the proof of Exercise 3.20.

Part 2: See the Hugin network. In the network for Exercise 3.19 we need to specify 90 numbers, but by decomposing the or-gate we only need 58 numbers.

Solution for exercise 3.22

See the Hugin network (there may be two many states for the free Hugin software to handle).

Solution for exercise 3.23

Solution for exercise 3.24

Solution for exercise 3.25

My computer did not allow more than five time slices.

Solution for exercise 3.26

The probability of a poor team winning the final is 0.043; instantiate the constraint node to '3' i the Hugin network. Note that in accordance with the prior probability of a team being poor, the T_i nodes are given the probabilities

$P(T_i) = (3/8, 5/8)$, however, due to the constraint node we would have gotten the same result if we had given the T_i nodes an even distribution..

Solution for exercise 3.27

(1) See the Hugin network. As the probability of $Result = y$ is positive, there are assigments of truth values making the expression true.

(ii) Insert $A = n$ and $B = n$ as evidence and propagate. As $P(Result = y) > 0$, there are assigments of the remaining variables making the expression true. If you insert "$Result = y$" and propagate, you see that the assigments must be $C = y, D = y, E = y, F = y$.

(iii) Assume that $\boldsymbol{E}$ is in conjunctive normal form. Construct a Bayesian network by introducing a variable for each Boolean variable, a variable for each conjunct, and a result variable. The parents of a conjunct variable are the Boolean variables in it, and the conjunct variables are related to the result variable through an "and"-relation. Give the Boolean variables a fifty-fifty prior. The resulting probability of $Result = y$ is positive if and only if there is at least one configuration over all variables (with $Result = y$) with a positive probability. Such a configuration is not inconsistent.

(iv) In iii) we have reduced the satisfiability problem to probability propagation in Bayesian networks. The BN-representation is polynomial in the length of the representation of the satisfiability problem.

Solution for exercise 3.28

See the Hugin network. By using this network we get $P(\text{Expert}) = (0.2, 0.8)$.

Solution for exercise 3.29

(i) The system (Hugin) responds with an error message "Inconsistence or overflow".

(ii) The posterior probability distributions are $P(Cold?|e) = (0.99, 0.01)$, $P(Angina?|e) = (0, 0.98, 0.02)$, $P(e) = 3.6 \times 10^{-6}$. Multiplying this with $P(Cold?|e)$ from above we get $P(Angina?, Cold? = no) = (0, 0.97, 0.02)$ $P(Angina?, Cold? = yes) = (0, 0.01, 0)$.

(iii) The posterior distributions after max-propagation are $Max(Cold?|e) = (1, 0.01)$ and $Max(Angina?|e) = (0, 1, 0.02)$. This shows that $(Cold? = no, Angina? = mild)$ is the most probable posterior configuration.

(iv) $Conf(e) = log(0.82 \times 0.99 \times 0.002) - log(3.6 \times 10^{-6}) = log428.45 = 8.7$. This indicates a strong conflict.

(v) Let $P(Angina? = mild, e) = x(s) = as + b$ and $P(e) = y(s) = cs + d$. Then $P(Angina?|e) = \frac{x(s)}{y(s)}$. To determine the constants we have from above: $0.01c + d = 3.6 \times 10^{-6}$, and through entering the virtual evidence $Angina? = mild$ we get $0.01a+b = 3.53\times10^{-6}$. We change the parameter s to 0.02, and we get $0.02c+d = 7.13\times10^{-6}$, and $0.02a+b = 7.05\times10^{-6}$. We now solve the two pairs of linear equations and get $a = 3.52\times10^{-4}, b = 10^{-8}, c = 3.53 \times 10^{-4}, d = 7 \times 10^{-8}$.

This yields $P(Angina? = mild|e) = \frac{3.52s+0.0001}{3.53s+0.0007}$.