Solutions for Bayesian networks and decision graphs (second edition)

Finn V. Jensen and Thomas D. Nielsen

May 19, 2008

Solution for exercise 4.1

$P(A = y|D = y) = 0.197$; $P(C = y|D = y) = 0.531$.

Solution for exercise 4.2

Solution for exercise 4.3

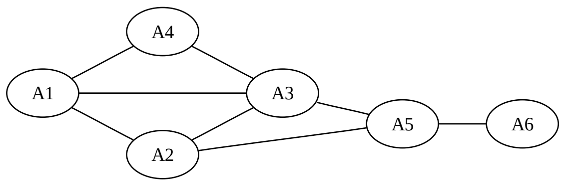
(i) $P(A = y|D = y) = 0.229$; $P(B = y|D = y) = 0.596$; $P(C|D = y) = 0.201$.
(ii) $P(B = y|C = y) = 0.201$.

Solution for exercise 4.4

(i) $3^6 = 729$.
(ii) $3^4 + 3^3 + 3^3 + 3^3 + 9 = 171$.
(iii) The elimination sequence F, D, E, B, C yields $4 \times 3^3 + 9 = 117$.

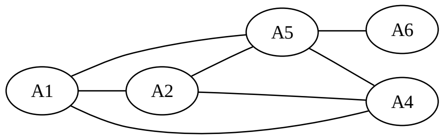
Solution for exercise 4.5

(i)



(ii) New links (fill-ins): $(A_1, A_5), (A_2, A_4), (A_4, A_5)$.

(iii)



Solution for exercise 4.6

Solution for exercise 4.7

Definition 4.1 is modified: line 2 is substituted by the line.
2'. Calculate $sx = |sp(\bigcup_{\phi \in \Phi_X} \mathrm{dom}\,(\phi))|$
Perform this operation in the prescribed elimination order, and collect the max of the sxs.

Solution for exercise 4.8

Solution for exercise 4.9

After elimination of a variable Y the new set of potentials represents $\Sigma_Y \Pi \Phi$:
$\Sigma_Y \Pi \Phi = \Sigma_Y [\Pi \Phi_Y \Pi_{\phi \notin \Phi_Y} \phi] = \Pi_{\phi \notin \Phi_Y} \phi \cdot \Sigma_Y \Pi \Phi_Y$, where $\Sigma_Y \Pi \Phi_Y = \Phi^{-Y}$.

Solution for exercise 4.10

Solution for exercise 4.11

The domains could be $[(A, B, E, F), (A, B, D, F), (A, B, C, D), (B, D, F, G)]$.
A (perfect) elimination sequence could be $< G, E, F, A, B, D >$.

Solution for exercise 4.12

Elimination order: $A_6, A_5, A_4, A_3, A_2, A_1$.

Solution for exercise 4.13

This domain graph has no perfect elimination sequence. After elimination of A_3, A_5, A_7, A_8, there no more simplicial nodes left.

Solution for exercise 4.14

ii. No. After elimination of H, G, F you are stuck.

Solution for exercise 4.15

(i) A and G.
(ii)No. After eliminating A and G, there are no simplicial nodes left.

Solution for exercise 4.16

Solution for exercise 4.17

For the sake of the exposition, a *large cycle* is a cycle of length over three.

(i) Induction. Certainly, any triangulated graph with less than four nodes is chord-saturated. Assume that all triangulated graphs with less than i nodes are chord-saturated, and let G be a triangulated graph with i nodes. Let X be a simplicial node in G. When eliminating X you get a triangulated graph with less than i nodes. The induction hypothesis yields that any large cycle in G not passing through X has a chord. Now consider a cycle passing through X. It has to pass two different neighbors Y and Z of X, and as X is simplicial, there is an edge connecting Y and Z. This edge is a chord in the cycle.

(ii) Let G be an incomplete chord-saturated graph with at least three nodes, and let A and B be non-adjacent. Let S be a minimal set of nodes (not containing A or B) such that any path connecting A and B contains a node from S. It is sufficient to prove that S is complete.

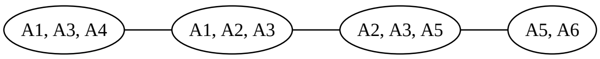
As S is minimal it must hold that for any $X \in S$ there must be a path from A to B, with X as the only member from S. Now, let $X, Y \in S$, and let p_X and p_Y be paths which contain only X and Y from S, respectively. Call the nodes between A and X, Y the $A-part$, and the nodes between B and X, Y the $B-part$. There must be a cycle containing X, Y and with a least one A-part node and at least one B-part node. Consider such a cycle with a minimal set of A-part and B-part nodes. As the A-part is minimal it cannot have a chord. The same holds for the B-part. If there is no edge between X and Y, then there must be a chord connecting an A-part node and a B-part node. Then there is a path connecting A and B without nodes from S. Therefore, X and Y must be adjacent.

(iii) Induction. Clearly, a chord-saturated graph with less than four nodes is triangulated. Assume that all chord-saturated graphs with less than k nodes is triangulated. Let G be a chord-saturated graph with k nodes. If G is complete, it is triangulated. If G is not complete, let A and B be non-adjacent. From (ii) we get a complete set S such that all paths from A to B pass through S. Consider G^A, the part of G consisting of nodes on paths between A and S (including A and S). It is sufficient to prove that G^A has a simplicial node not in S, as we can then eliminate the simplicial node and obtain a cord-saturated graph with less than k nodes. Clearly G^A is chord-saturated, and by the induction hypothesis, G^A is triangulated. From the proof of Theorem 4.1 we have that G^A has two non-adjacent simplicial nodes, hence at least one of them cannot be in S.

Solution for exercise 4.18

Solution for exercise 4.19

(i) The cliques are exactly the domains of the potentials.
(ii)

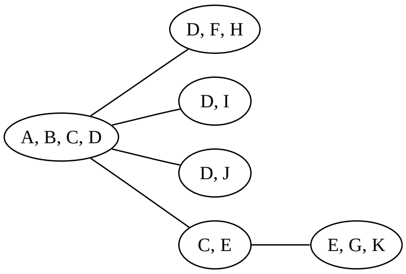


Solution for exercise 4.20

(i) $(A, B, C, D), (A, B, D, E), (B, C, D, F)$.
(ii) $(A, B, D, E) - (A, B, C, D) - (B, C, D, F)$.

Solution for exercise 4.21

The cliques are $(A, B, C, D), (D, F, H), (D, I), (D, J), (C, E), (E, G, K)$. This is exactly the domains of the potentials of the non-top variables.



Solution for exercise 4.22

Solution for exercise 4.23 (i) Let MG be the moral graph for G, and let T be the graph resulting from removing the directions from the links in G. As T is a tree there will always be at least one leaf L. A leaf is characterized by having only one neighbor. If the neighbor is a parent, elimination of L will not add any links in MG, and the graph resulting from removing L from T is still a tree. If the neighbor is a child C, neighbors of L in MG are the parents of C. Therefore the neighbor set of L is complete in MG, and L is simplicial in MG.

(ii)**claim** If neighbor A and B share the neighbors C and D then C and D are neighbors.
proof of claim: Assume that C and D are not neighbors. Then they cannot be parents together for neither A nor B nor can they both be parents together with A or B. For any other possible scenario (C parent of A and B together with D a common child of A and B; C and D both children of A and B, ...), the graph is not singly connected.

proof of (ii): Let C_1 and C_2 be two cliques sharing A, B; let $C \in C_1$ and $D \in C_2$. The claim yields that $\{A, B, C, D\}$, and ultimately $C_1 = C_2$.

Solution for exercise 4.24

Take the order by collecting to (A, B, C, D) and afterwards distributing from it. In the collect phase all messages are unit potentials ($\mathbf{1}$) except: (D, I) transmits $P(i|D)$ and (C, E) transmits $P(e|C)$ to (A, B, C, D). In the distribute phase only (A, B, C, D) transmits non-constant potentials:
$\{\Sigma_{A,B,C} P(D|A, B, C)P(A)P(B)P(C), P(i|D)\}$ to (F, D, H) and to (D, J),
$\{\Sigma_{A,B,C} P(D|A, B, C)P(A)P(B)P(C)\}$ to (D, I) (which is actually pointless), and $\Sigma_{A,B,D} P(D|A, B, C)P(A)P(A)P(B)P(i|D), P(C)\}$ to (C, E).

Solution for exercise 4.25

Unfortunately, this exercise is not well formulated.

We assume that we have the following potentials: $\phi_1(A,B,C,D), \phi_3(D,E,F,I)$, $\phi_5(C,G,H,J), \phi_6(B,C,D,G), \phi_{10}(B,C,D,E)$.

Cliques V_i are attached potentials ϕ_i

The formulation is not clear on what to collect to. If you (rather unconventionally) collect to the *separator* CG you have the following messages:

From V_5: $\Sigma_{HJ}\phi_5$.
From V_1 to S_1 and V_6: $\Sigma_A\phi_1$
From V_3 to S_3 and V_{10}: $\Sigma_{FI}\phi_3$
From V_{10} to S_6 and V_6: $\Sigma_E[\phi_{10}\Sigma_{FI}\phi_3]$
From V_6 to S_5: $\Sigma_{BD}[\phi_6\Sigma_E(\phi_{10}\Sigma_{FI}\phi_3)\Sigma_A\phi_1]$

Solution for exercise 4.26

i. Links: $(A,B,C) - [C] - (C,D,E), (C,D,E) - [D,E] - (D,E,G)$, $(C.D.E) - [C,E] - (C,E,F), (C,E,F) - [E,F] - (E,F,H)$.

Solution for exercise 4.27 Assume that the direction $V_i \rightarrow V_j$ is not triggered (or used). Then at least one direction $V_k \rightarrow V_j$ $(k \neq j)$ is not used. If it is triggered, then a message can be sent. If not triggered, then continue the reasoning. Eventually you reach a leaf. A leaf is always triggered, and a message can be sent.

Solution for exercise 4.28

Let R be the first node to receive all of its messages. Then all messages passed so far must have been sent in the direction of R: suppose not, and assume that V has sent a message away from R. At that instance V must have received a message from the R-direction. When R has received all messages, it must also have received a message from the V-direction. This means that prior to this, has sent a message in the R-direction, and that time V has received all its messages, which is before R has received all its messages.
Hence, the transmission until R has received all its messages corresponds to a CollectEvidence to R. The remaining messages correspond to a DistributeEvidence from R.

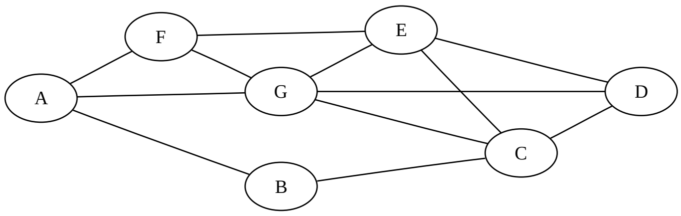
Solution for exercise 4.29

Add a fill-in (A_1, A_4) or (A_2, A_6). The first is the result of the elimination sequence A_3, A_5, A_7, A_8, A_2, and the latter is the result of the sequence $A_3, A_5, A_7, A_8, A_1, ...$

Solution for exercise 4.30

ii. The elimination sequence is $F, J, B, D, A, E, C, ...$

Solution for exercise 4.31



An elimination sequence could be D, B, C, A, F, G, E. The result would be the same if all nodes had two states. Actually, if all nodes have the same number states, then this heuristic methods boils down to counting neighbors.

Solution for exercise 4.32

A better elimination sequence is $A, C, H, I, J, D, E, ...$

Solution for exercise 4.33

(i)Since $P(A, B, C, D) = P(A)P(B|A)P(C|A)P(D|B, C)$ we have
$P(B, C, D|A) = P(B|A)P(C|A)P(D|B, C)$, and therefore
$P(B, C, D|a) = P(B|a)P(C|a)P(D|B, C)$ (*) This yields that the model (b) in the Figure is a model for $P(B, C, D|a)$.

(ii) A result of (*) above.

(iii) Initially we have $P(A)$. The propagations yield $P(B, e|A)$. Then $P(B, e) = \sum_A P(b, e|A)P(A)$.
For each propagation, the normalization constant is $P(e, a)$. Hence, the propagations yield $P(e|A)$, and a multiplication with $P(A)$ will give $P(A, e)$

iv) Using Bayes' rule on A and B we get
$P(A, B, C, D) = P(B)P(A|B)P(C|A)P(D|B, C)$ and hence
$P(A, C, D|B) = P(A|B)P(C|A)P(D|B, C)$.
Therefore, $P(A, C, D|b) = P(A|b)P(C|A)P(D|b, C)$, and the model (c) in the

Figure is a model for $P(A, C, D|b)$.

v) Bayes' inversion on D as used in iv) requires $P(B, C)$.

vi) Conditioning on the variables A, and E yield a singly connected network (A, E is called a *minimal cut set*).

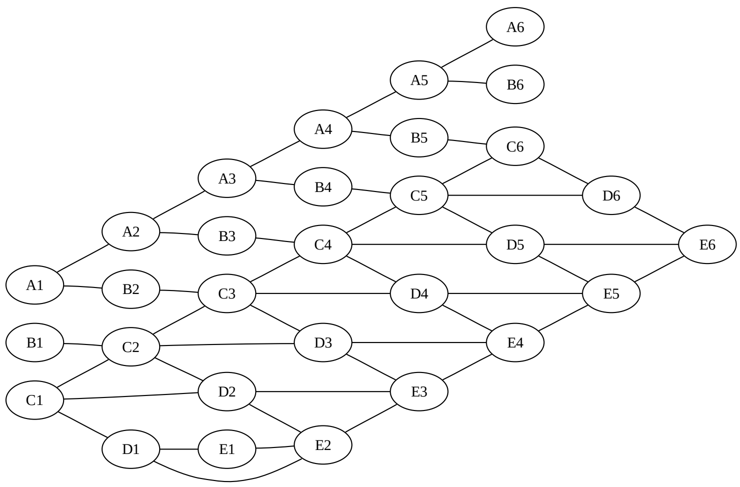
vii) Conditioning on $A$and E results in 10 networks, while conditioning on A, B and G results in 8 networks only.

Solution for exercise 4.34

Solution for exercise 4.35

(i) First eliminate the simplicial nodes $A_4, B_4, D_4, A_3, B_1, E_1$; eliminating B_3, B_2 will not really increase complexity; then A_1, A_2, C_4, C_3 can be eliminated as simplicial; when eliminating the rest in the order $D_3, D_1, C_1, E_3, E_2, D_2, C_2$, no clique will contain more than four nodes.

(ii)



Following a similar strategy as above, we eliminate the simplicial nodes $A_6, B_6, D_6, E_6, B_1, E_1$; eliminating B_2, B_3, B_4, B_5 followed by D_2, A_1, A_4, A_3 will not really increase complexity. But now all nodes have at least three neighbors, and it is not possible to continue without at some stage to eliminate a node with four neighbors. **Note**: this is not a proof that it can't be done.

Solution for exercise 4.36

Solution for exercise 4.37

Solution for exercise 4.38

Solution for exercise 4.39

Marginals for the state y from Table 4.2; $A : 0, 39, B : 0.62, C : 0.51, D : 0.32, E : 0.99$
Marginals from the model (Table 4.1): $A : 0, , B : 0.6, C : 0.52, D : 0.34, E : 0.98$

4

Solution for exercise 4.40

Solution for exercise 4.41

NA

Solution for exercise 4.42

Solution for exercise 4.43

For this network it holds that when the of two variables is known, then the state of the third variable is also known. Therefore the Gibbs sampling will never leave its start configuration.

Solution for exercise 4.44