

Bayesian Networks

6. Exact Inference / Examples

Lars Schmidt-Thieme

Information Systems and Machine Learning Lab (ISMLL)
Institute for Business Economics and Information Systems
& Institute for Computer Science
University of Hildesheim
<http://www.isml.uni-hildesheim.de>

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), Institute BW/WI & Institute for Computer Science, University of Hildesheim
Course on Bayesian Networks, summer term 2010

1/33

Bayesian Networks



1. Studfarm

2. Hailfinder

Studfarm / problem

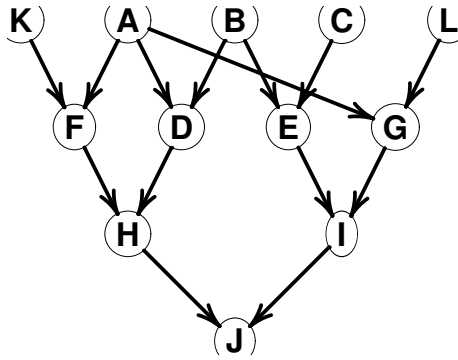


Figure 1: Studfarm bayesian network. potentials:

(i) $p(X)$ for $X = A, B, C, K, L$:

$$p(X = aa) = 0.99, \quad p(X = aA) = 0.01$$

(ii) $p(X|Y, Z)$ for $(X|Y, Z) = (D|A, B), (E|B, C), (F|A, K), (G|A, L), (H|F, D), (I|E, G)$:

father Y	aa		aA	
mother Z	aa	aA	aa	aA
aa	1	.5	.5	.25
aA	0	.5	.5	.25

(iii) and $p(J|H, I)$:

father H	aa		aA	
mother I	aa	aA	aa	aA
aa	1	.5	.5	.25
aA	0	.5	.5	.5
AA	0	0	0	.25

Evidence is given, that John is sick, represented by $p(J)$:

$$p(J = AA) = 1$$

Studfarm / markov network

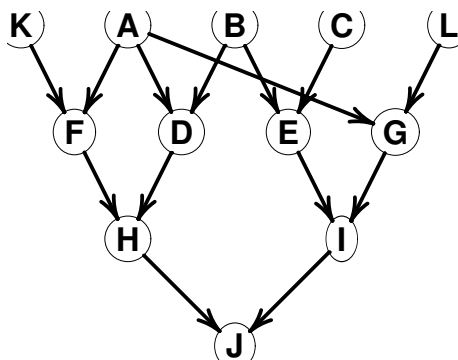


Figure 2: Studfarm bayesian network.

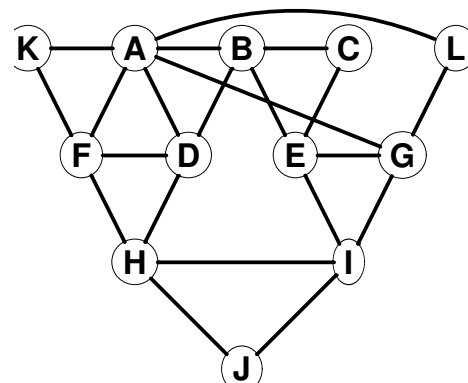


Figure 3: Studfarm markov network (moral graph).

Studfarm / triangulation & cluster tree (MCS)

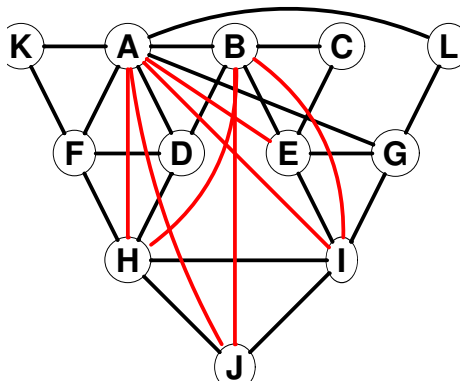


Figure 4: Triangulation of Studfarm markov network by MCS (fill-in 7).

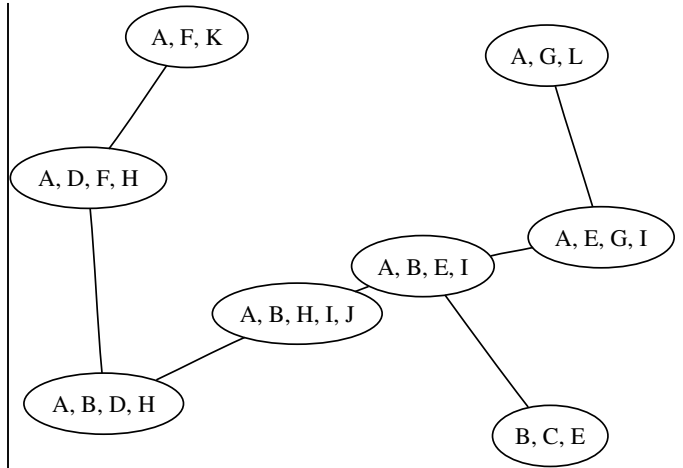


Figure 5: Cluster tree for the triangulation at the left (total state space size 136).

Studfarm / triangulation & cluster tree (Minimal degree)

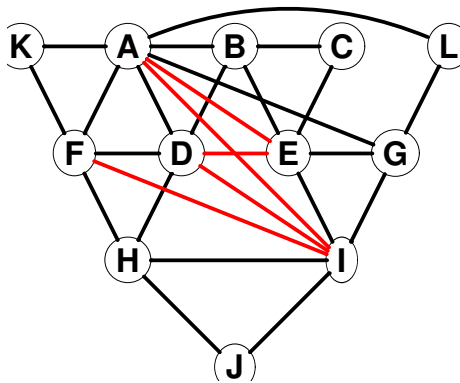


Figure 6: Triangulation of Studfarm markov network by Minimal Degree Heuristics (fill-in 5).

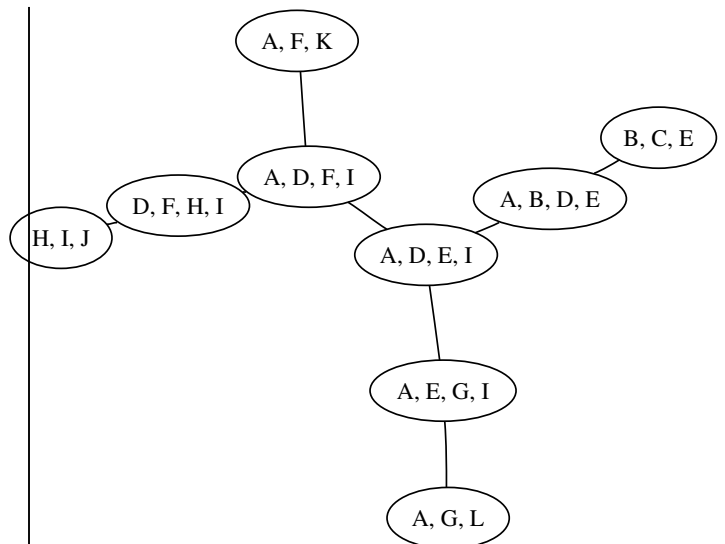


Figure 7: Cluster tree for the triangulation at the left (total state space size 116).

Studfarm / attach potentials

- (i) $p(A), p(B), p(C), p(K), p(L),$
- (ii) $p(D|A, B), p(E|B, C), p(F|A, K),$
 $p(G|A, L), p(H|F, D), p(I|E, G),$
- (iii) and $p(J|H, I),$
- (iv) and evidence $p(J).$

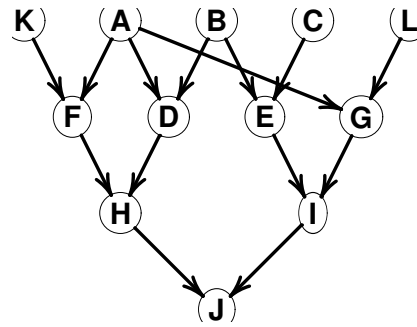
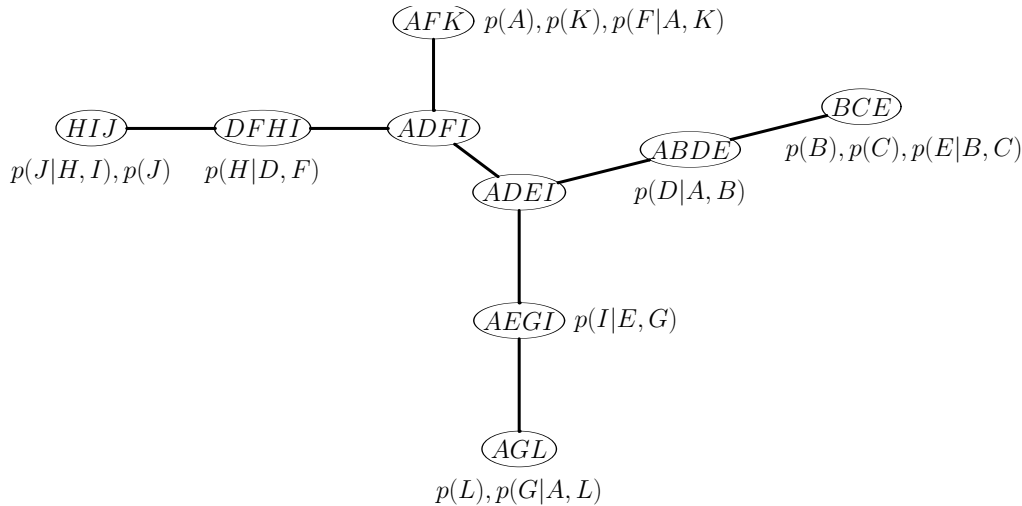


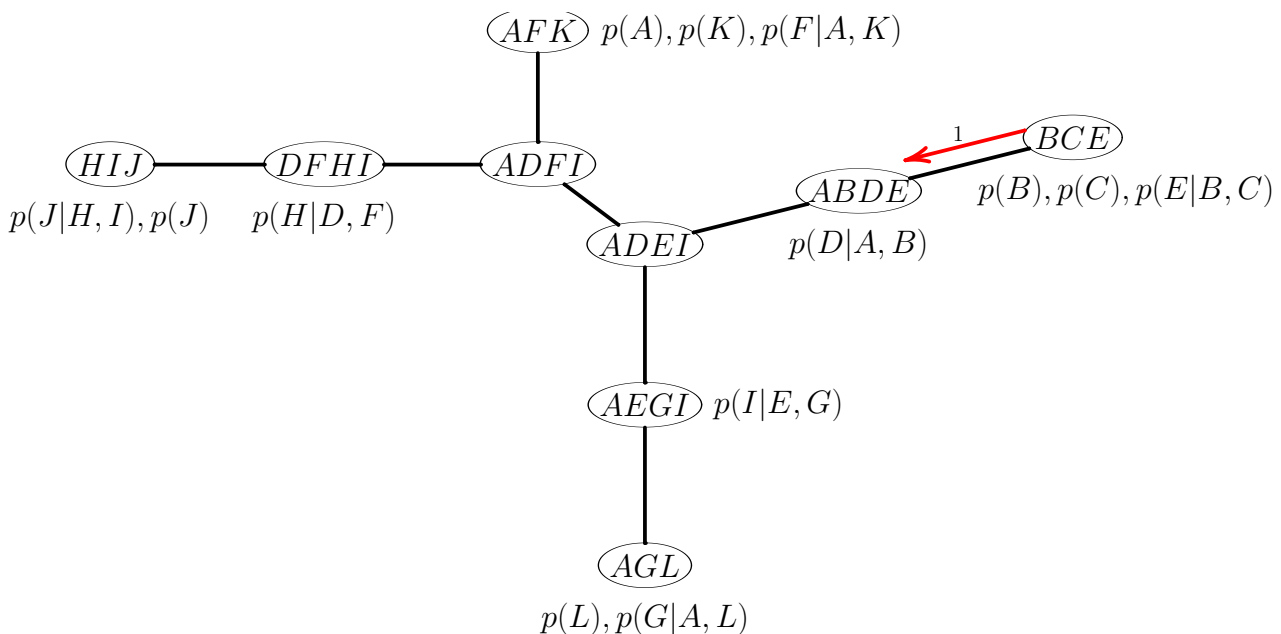
Figure 8: Studfarm bayesian network.



Studfarm / collect evidence (1/8)

collect evidence:

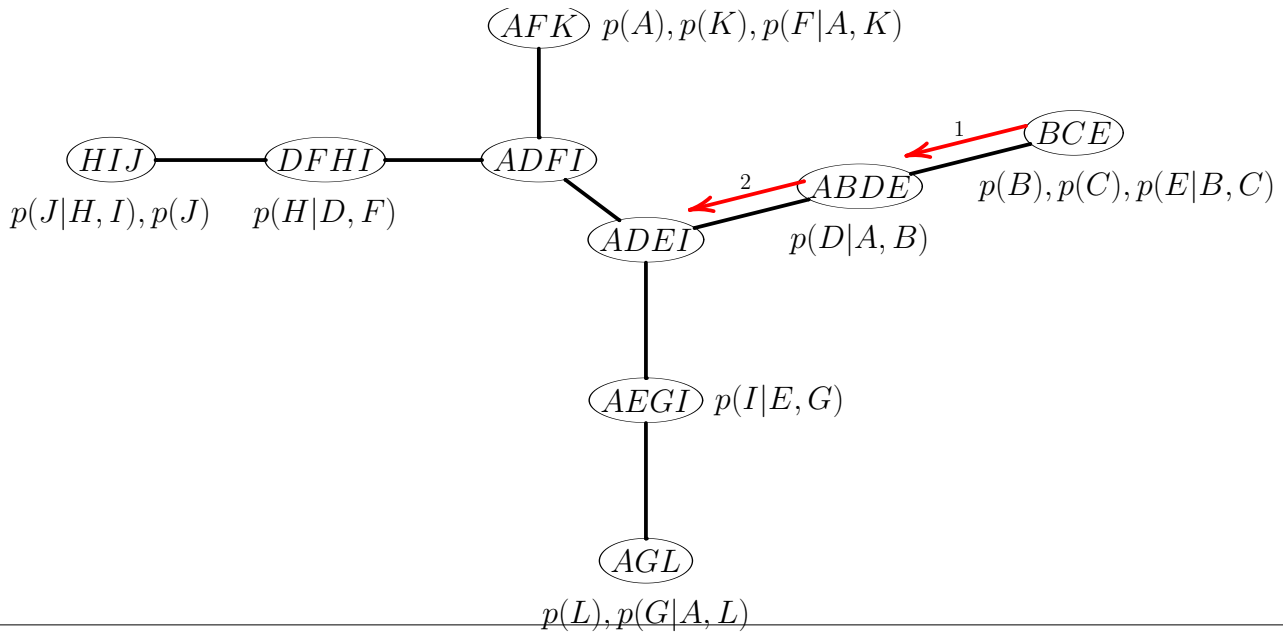
$$q_{BCE \rightarrow ABDE} := (p(B) \cdot p(C) \cdot p(E|B, C))^{\downarrow BE} = \begin{array}{c|cc} E & \text{pure carrier} & \\ \hline B = \text{pure} & 0.985 & 0.005 \\ \text{carrier} & 0.005 & 0.005 \end{array}$$



Studfarm / collect evidence (2/8)

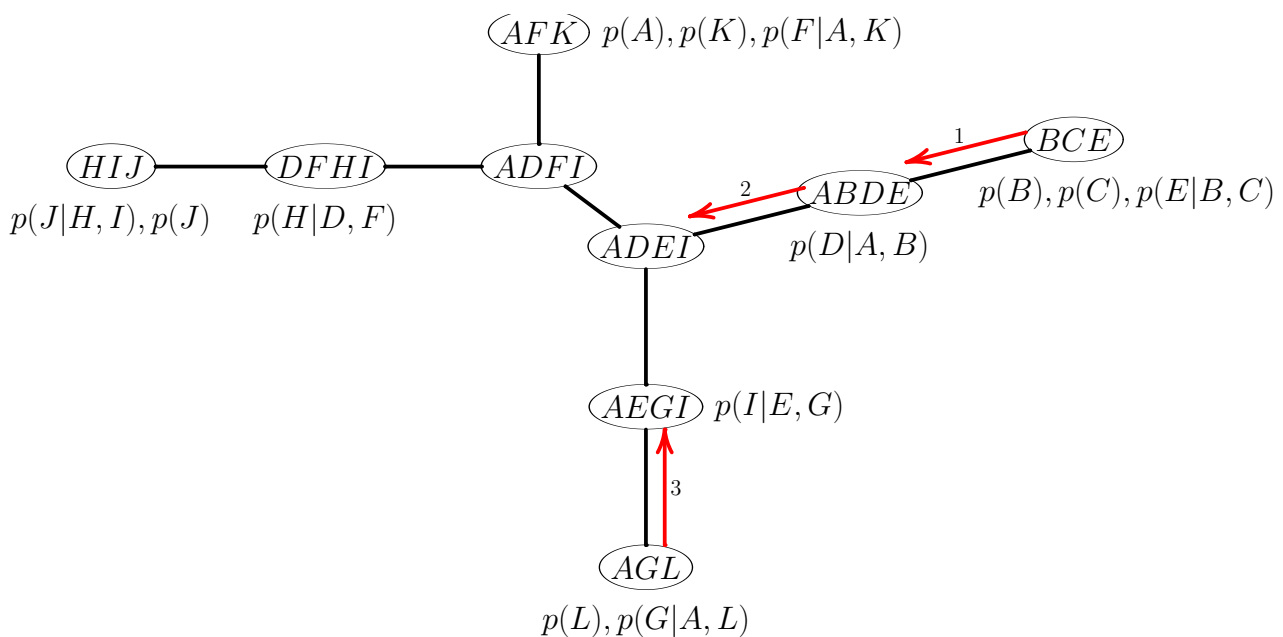
$$q_{ABDE \rightarrow ADEI} := (q_{BCE \rightarrow ABDE} \cdot p(D|A, B))^{\downarrow ADE}$$

	D pure		carrier	
	pure	carrier	pure	carrier
A = pure	0.9875	0.0075	0.0025	0.0025
carrier	0.4942	0.0041	0.4958	0.0058



Studfarm / collect evidence (3/8)

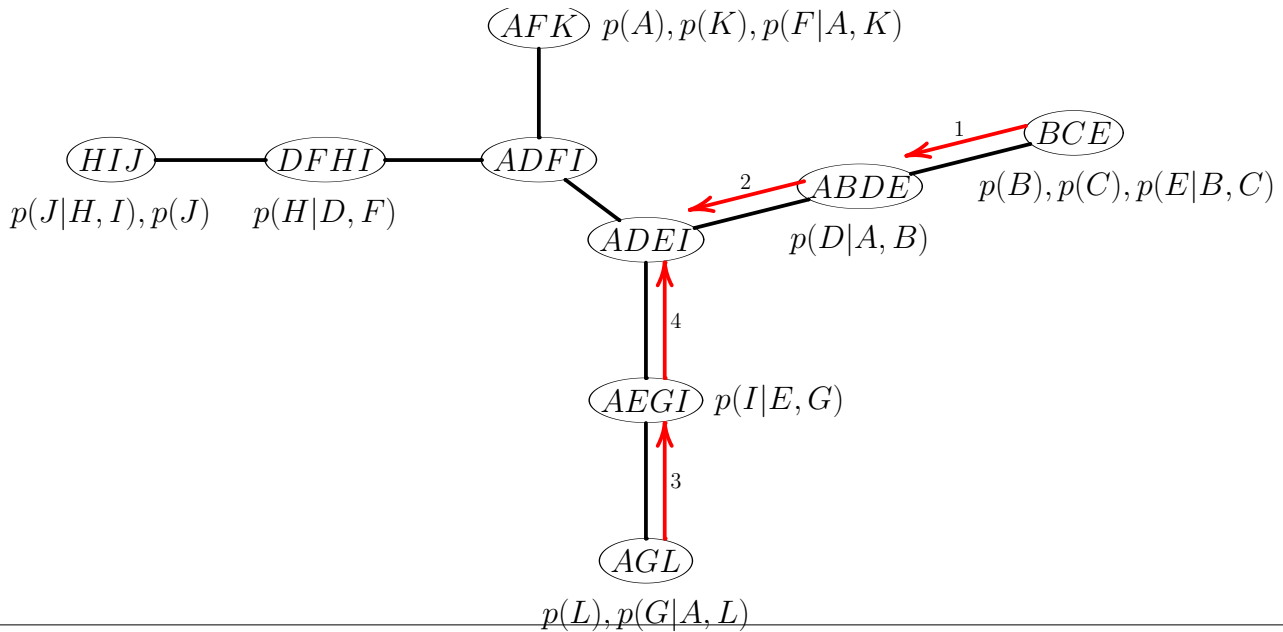
$$q_{AGL \rightarrow AEGI} := (p(C) \cdot p(G|A, L))^{\downarrow AG} = \begin{array}{c|cc} G & \text{pure} & \text{carrier} \\ \hline A = \text{pure} & 0.995 & 0.005 \\ \text{carrier} & 0.4983 & 0.5017 \end{array}$$



Studfarm / collect evidence (4/8)

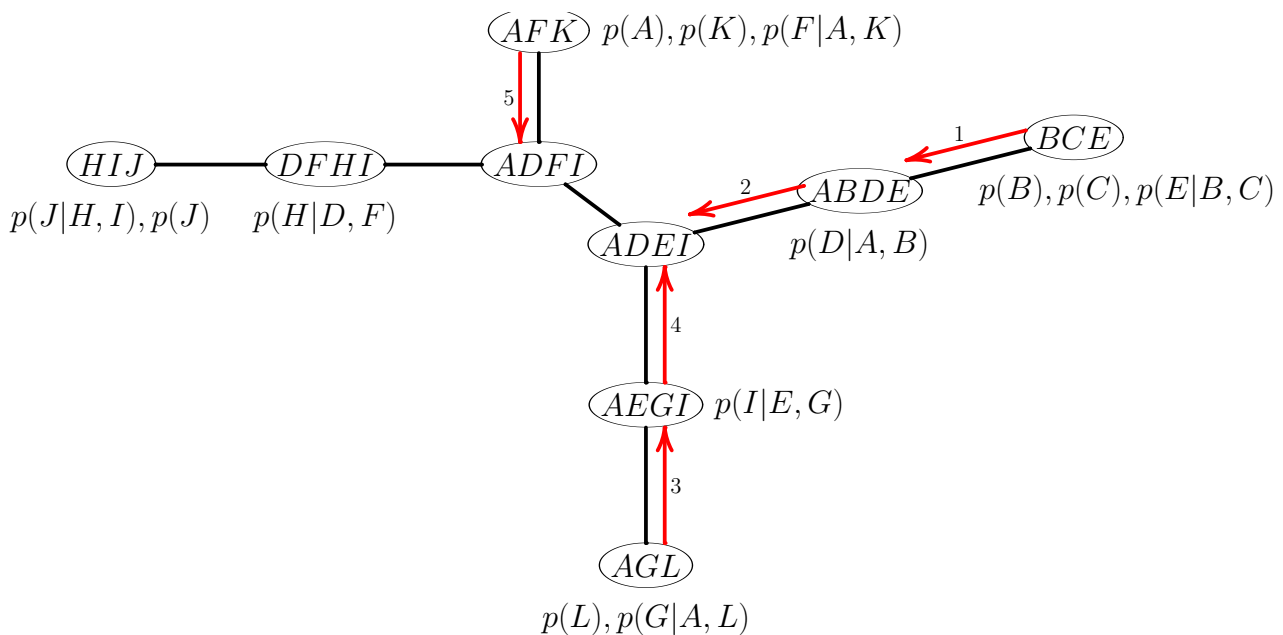
$$q_{AEGI \rightarrow ADEI} := (q_{AGL \rightarrow AEGI} \cdot p(I|E, G))^{\downarrow AEI}$$

	E pure		carrier	
	I pure	carrier		
A = pure	0.9975	0.0025	0.4992	0.5008
carrier	0.7492	0.2508	0.4164	0.5836



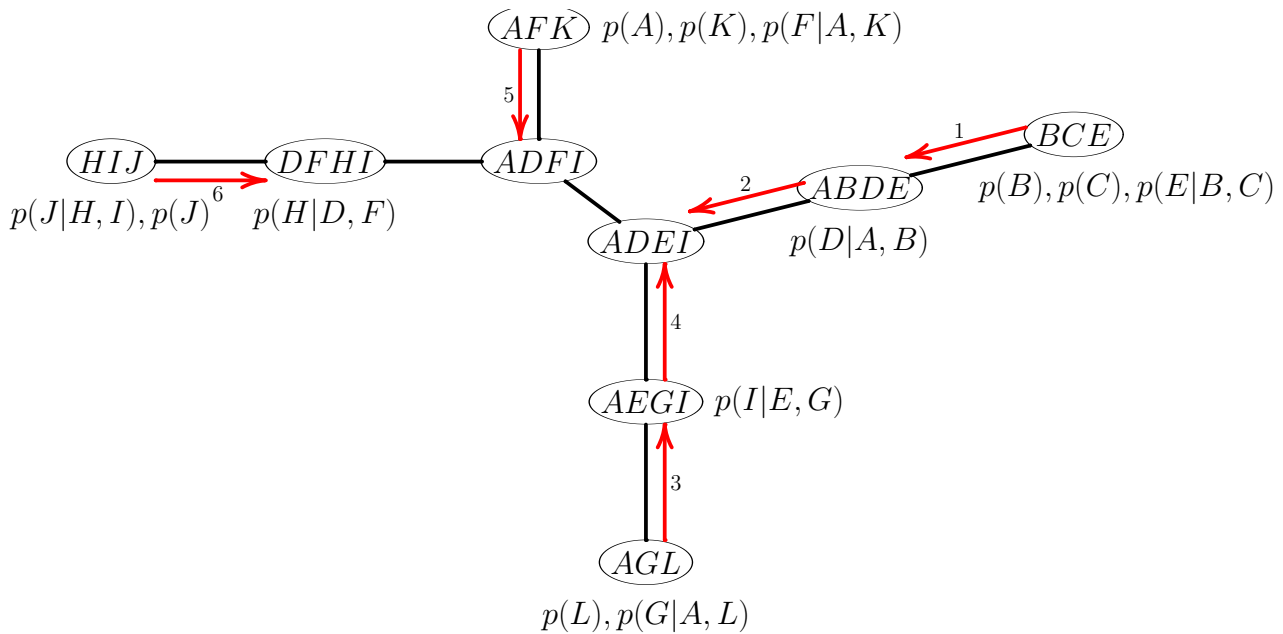
Studfarm / collect evidence (5/8)

$$q_{AFK \rightarrow ADFI} := (p(A) \cdot p(K) \cdot p(F|A, K))^{\downarrow AF} = \begin{matrix} F & \text{pure} & \text{carrier} \\ \hline A = \text{pure} & 0.985 & 0.005 \\ \text{carrier} & 0.005 & 0.005 \end{matrix}$$



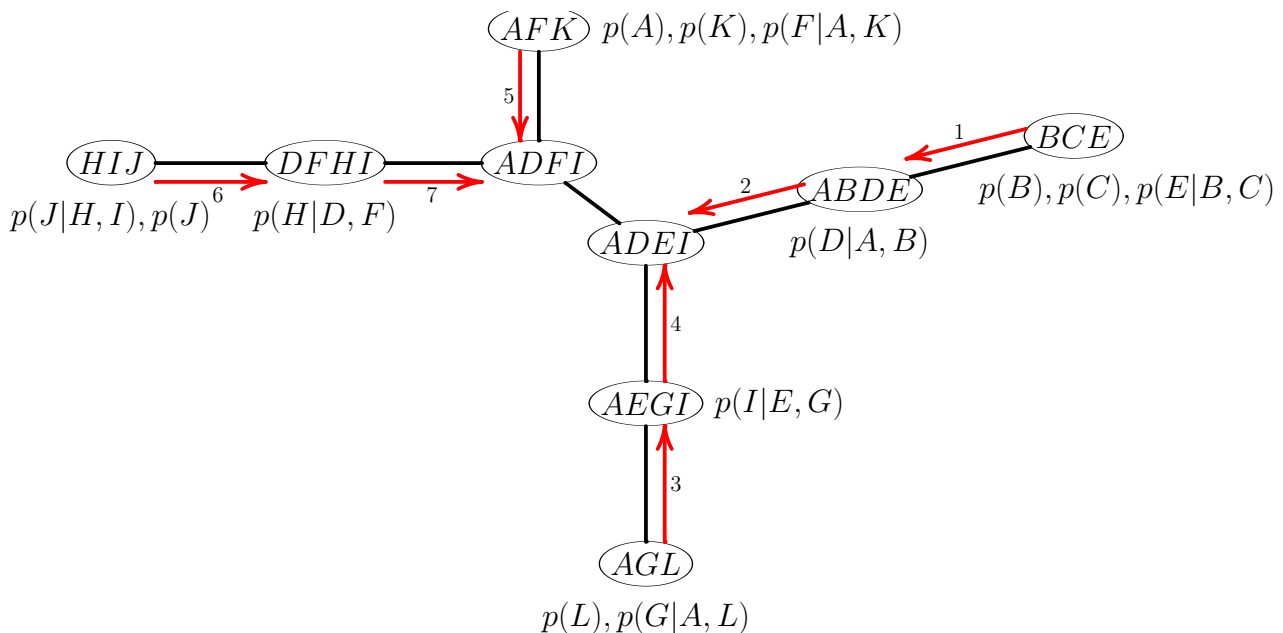
Studfarm / collect evidence (6/8)

$$q_{HIJ \rightarrow DFHI} := (p(J|H, I) \cdot p(J))^{\downarrow HI} = \begin{array}{c|cc} & I & \text{pure carrier} \\ \hline H = \text{pure carrier} & 0 & 0 \\ \hline \text{carrier} & 0 & 0.25 \end{array}$$



Studfarm / collect evidence (7/8)

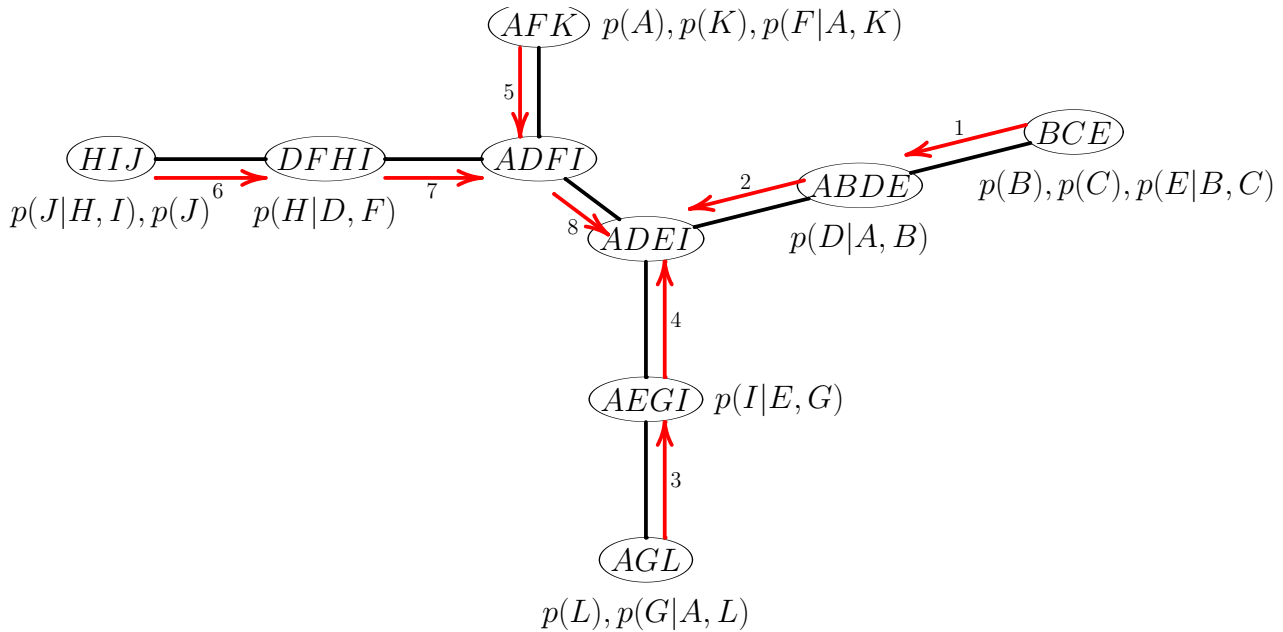
$$q_{DFHI \rightarrow ADFI} := (q_{HIJ \rightarrow DFHI} \cdot p(H|D, F))^{\downarrow DFI} = \begin{array}{c|ccc} & F & \text{pure} & \text{carrier} \\ & I & \text{pure} & \text{carrier} \\ \hline D = \text{pure} & 0 & 0 & 0 & 0.125 \\ \hline \text{carrier} & 0 & 0.125 & 0 & 0.1667 \end{array}$$



Studfarm / collect evidence (8/8)

$$q_{ADFI \rightarrow ADEI} := (q_{DFHI \rightarrow ADFI} \cdot q_{AFK \rightarrow ADFI}) \downarrow^{ADI} =$$

	D		carrier	
	pure	carrier		
I				
A = pure	0	0.0006	0	0.124
carrier	0	0.0006	0	0.0015

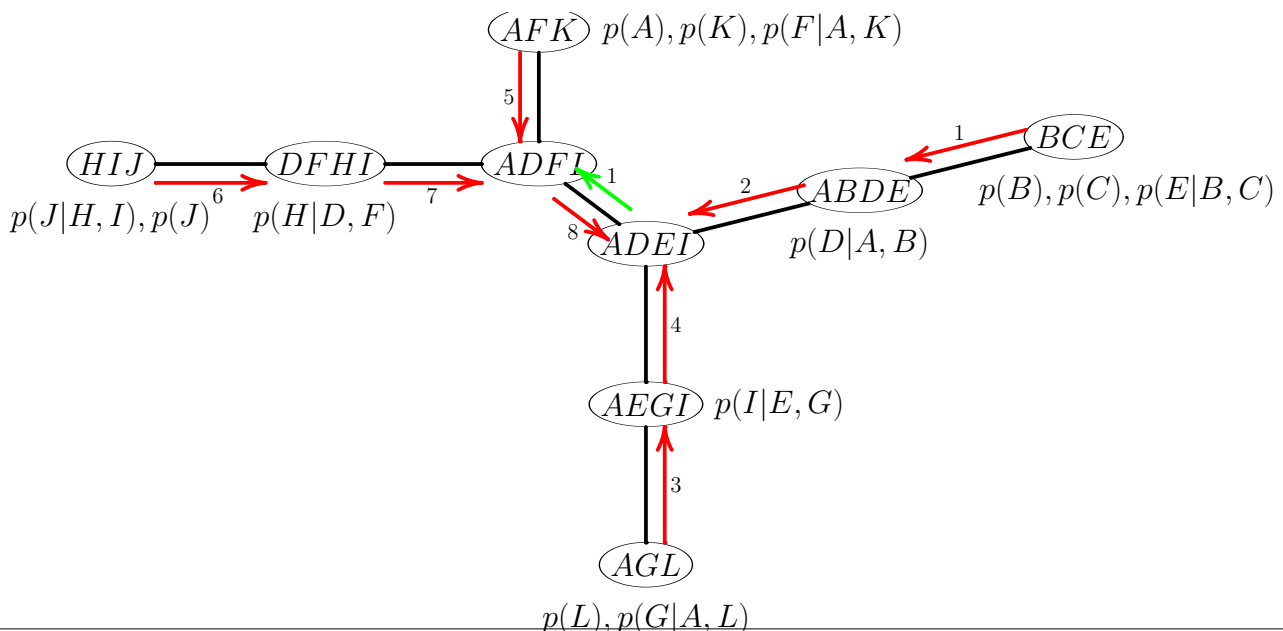


Studfarm / distribute evidence (1/8)

distribute evidence:

$$q_{ADEI \rightarrow ADFI} := (q_{ABDE \rightarrow ADEI} \cdot q_{AEGI \rightarrow ADEI}) \downarrow^{ADI}$$

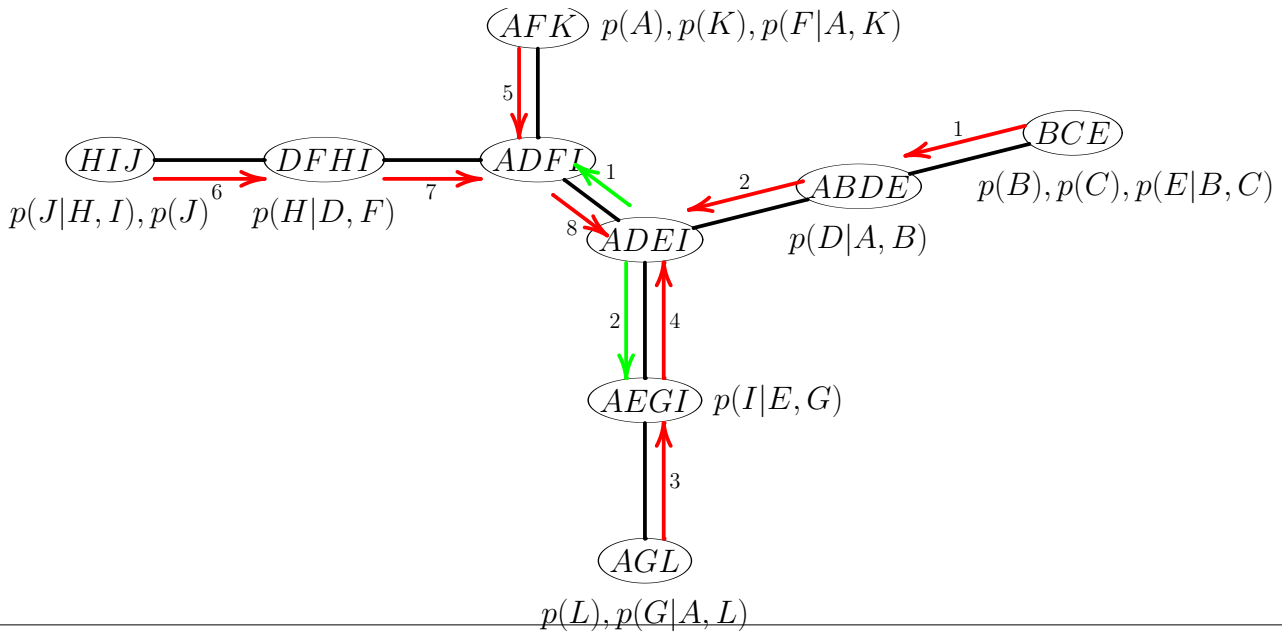
	D		carrier	
	pure	carrier		
I				
A = pure	0.9888	0.0062	0.0037	0.0013
carrier	0.372	0.1264	0.3739	0.1278



Studfarm / distribute evidence (2/8)

$$q_{ADEI \rightarrow AEGI} := (q_{ABDE \rightarrow ADEI} \cdot q_{ADFI \rightarrow ADEI}) \downarrow^{AEI}$$

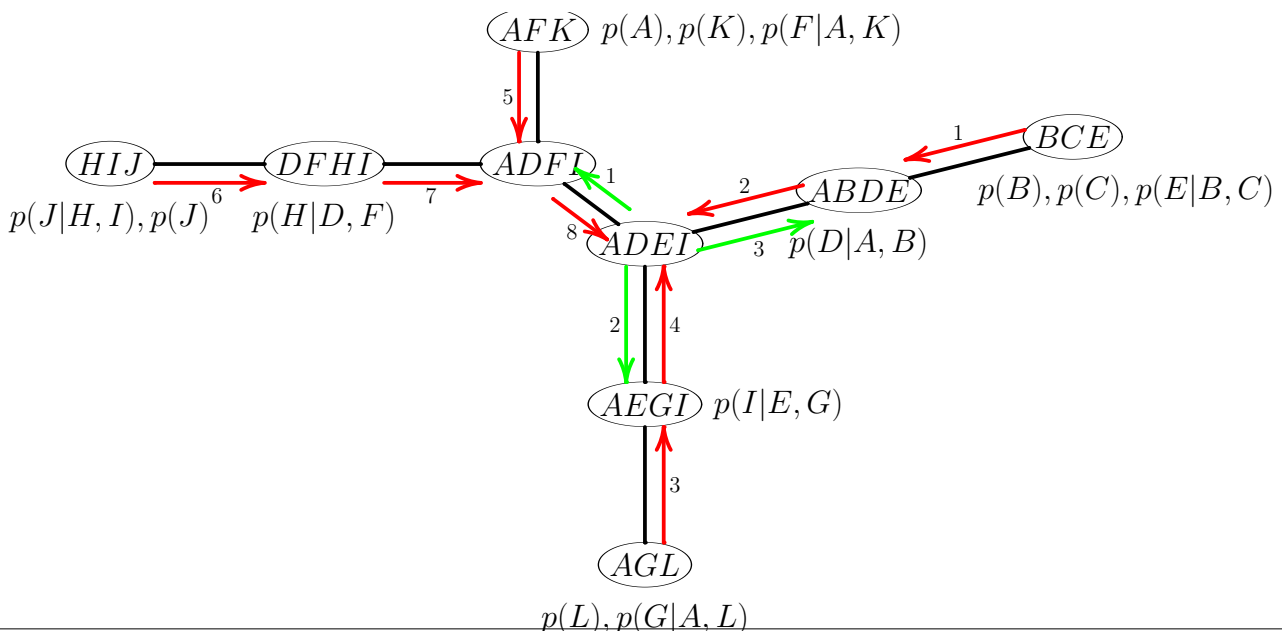
		E		I	
		pure	carrier	pure	carrier
A =	pure	0	0.0009	0	0.0003
	carrier	0	0.001	0	0



Studfarm / distribute evidence (3/8)

$$q_{ADEI \rightarrow ABDE} := (q_{AEGI \rightarrow ADEI} \cdot q_{ADFI \rightarrow ADEI}) \downarrow^{ADE}$$

		D		E	
		pure	carrier	pure	carrier
A =	pure	0	0.0003	0.0003	0.0621
	carrier	0.0002	0.0004	0.0004	0.0009

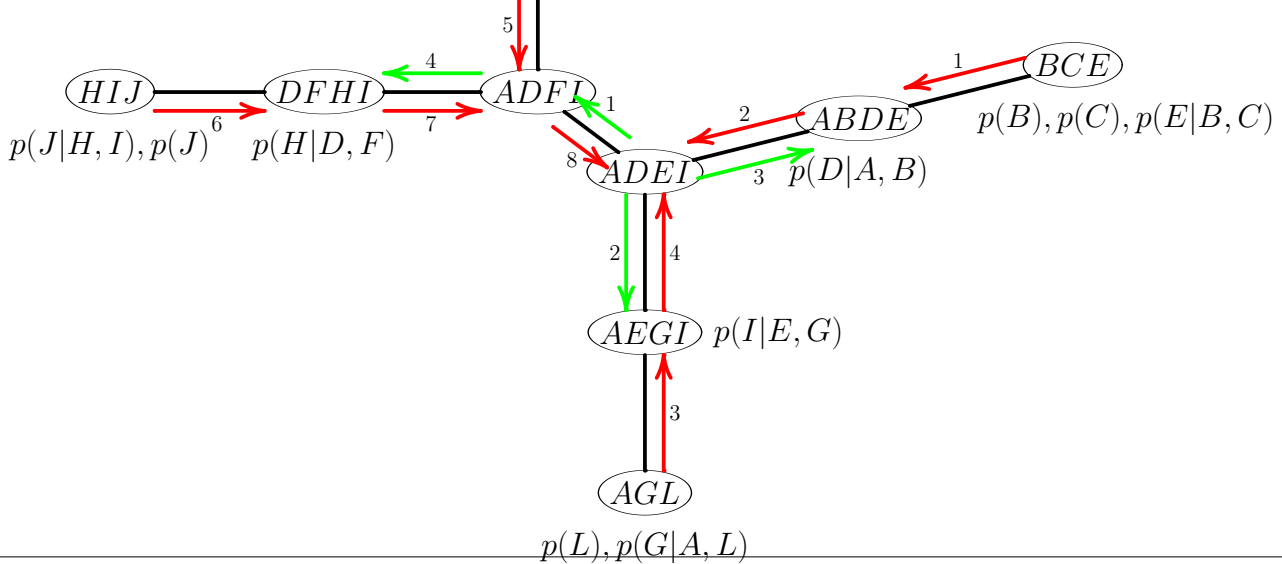


Studfarm / distribute evidence (4/8)

$$q_{ADFI \rightarrow DFHI} := (q_{ADEI \rightarrow ADFI} \cdot q_{AFK \rightarrow ADFI}) \downarrow^{DFI}$$

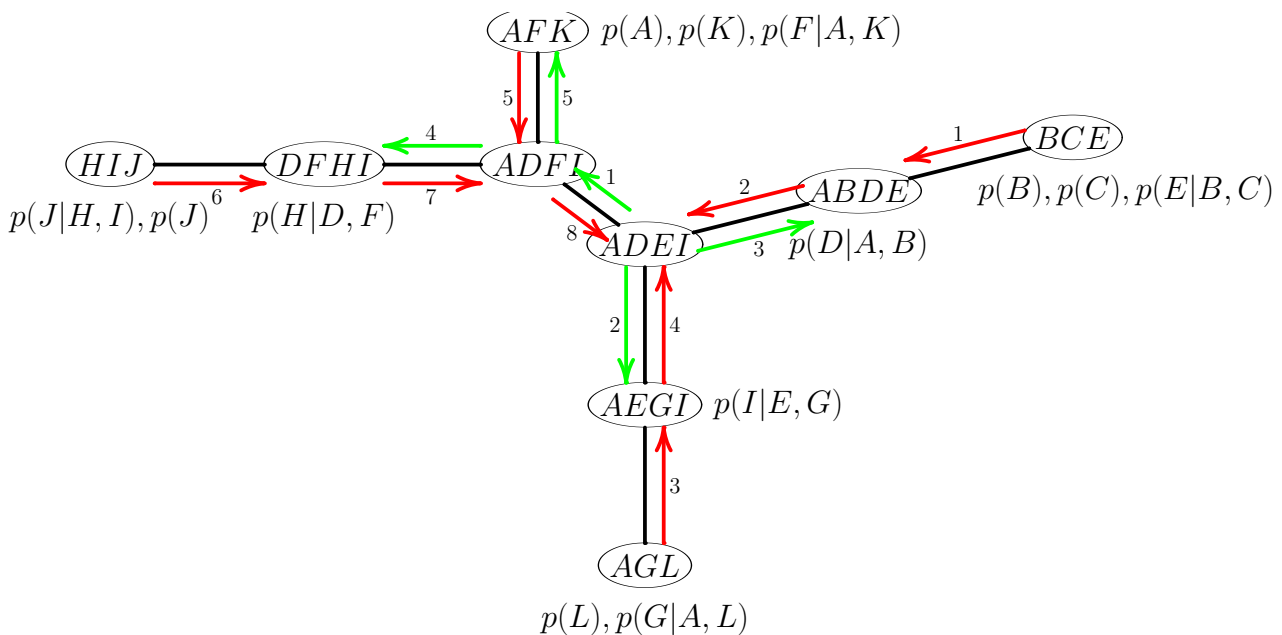
	F		carrier	
I	pure	carrier		
D = pure	0.9759	0.0067	0.0068	0.0007
carrier	0.0055	0.0019	0.0019	0.0006

(AFK) $p(A), p(K), p(F|A, K)$



Studfarm / distribute evidence (5/8)

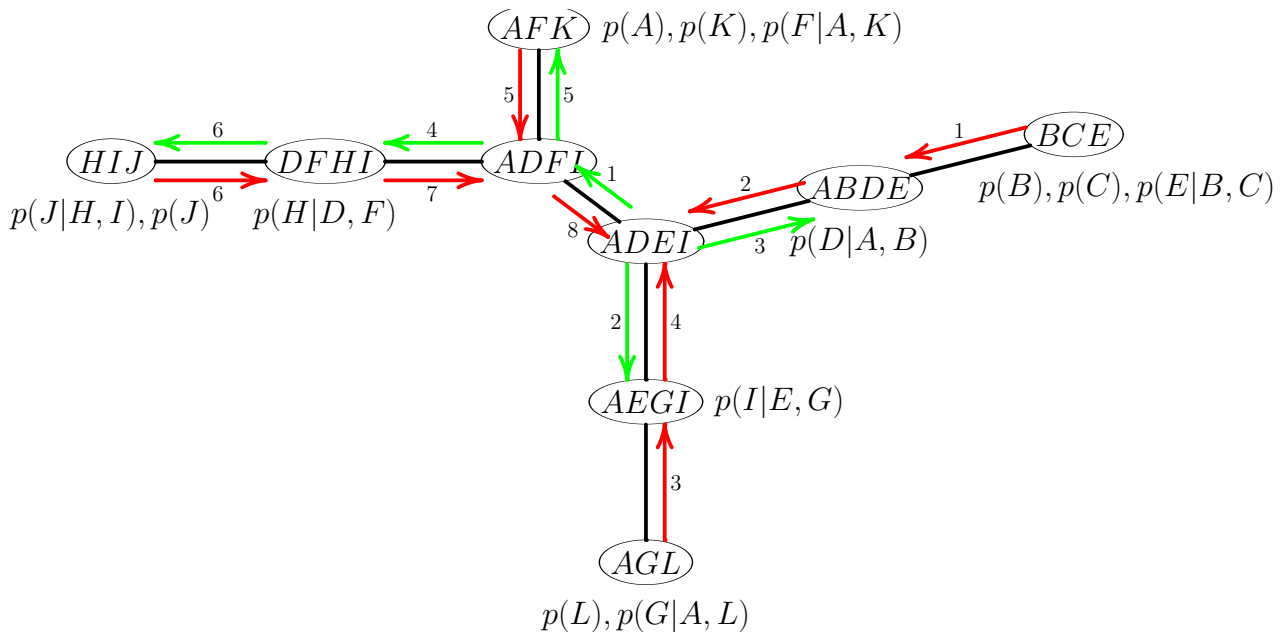
$$q_{ADFI \rightarrow AFK} := (q_{ADEI \rightarrow ADFI} \cdot q_{DFHI \rightarrow ADFI}) \downarrow^{AF} = \begin{array}{c|cc} & F & \text{pure carrier} \\ \hline A = \text{pure} & 0.0002 & 0.001 \\ \text{carrier} & 0.016 & 0.0371 \end{array}$$



Studfarm / distribute evidence (6/8)

$$q_{DFHI \rightarrow HIJ} := (q_{ADFI \rightarrow DFHI} \cdot p(H|D, F))^{\downarrow HI} =$$

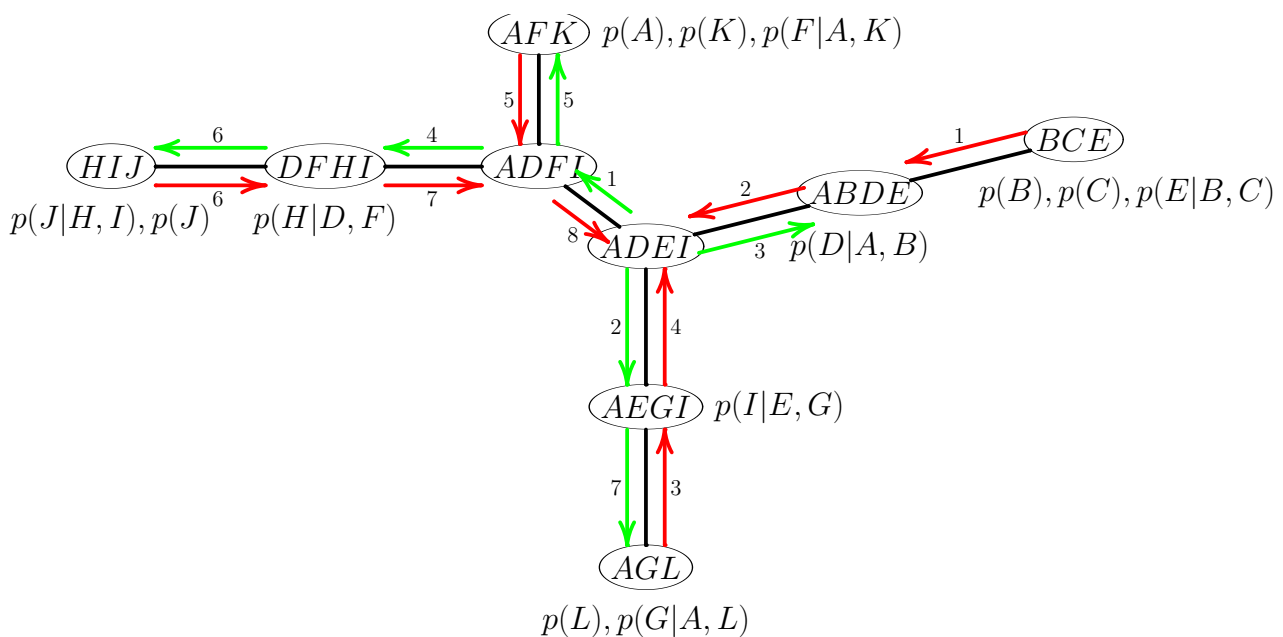
I	pure carrier	
$H = \text{pure carrier}$	0.9827	0.0082
carrier	0.0074	0.0017



Studfarm / distribute evidence (7/8)

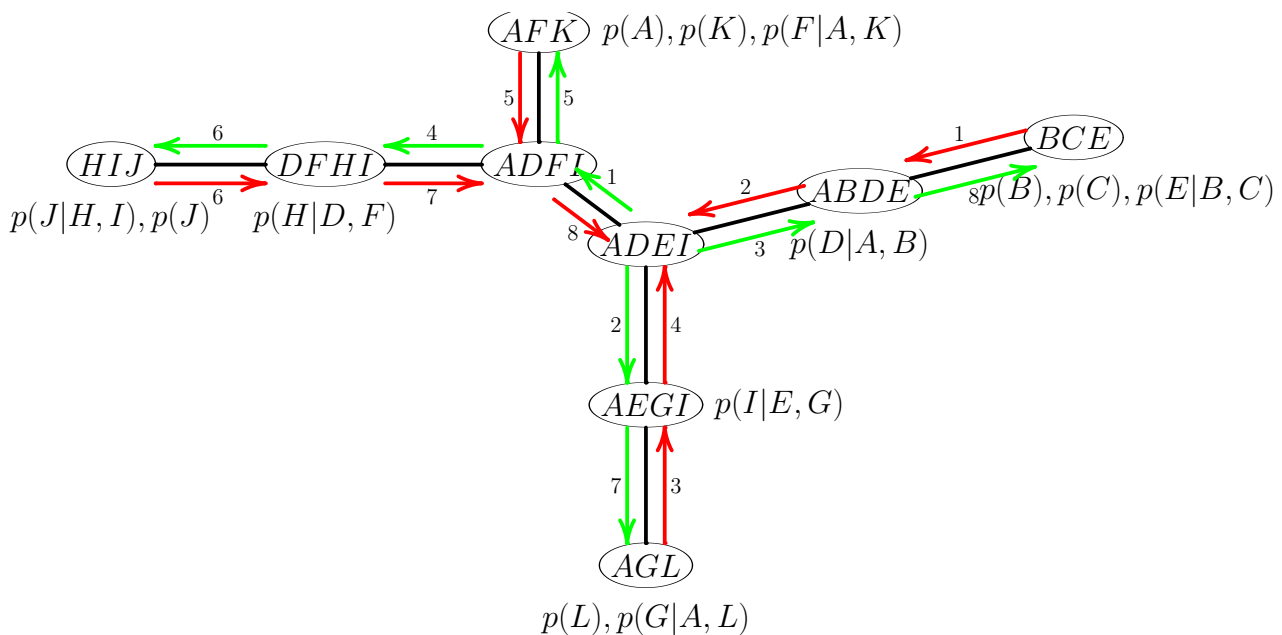
$$q_{AEGI \rightarrow AGL} := (q_{ADEFI \rightarrow AEGI} \cdot p(I|E, G))^{\downarrow AG} =$$

G	pure carrier	
$A = \text{pure carrier}$	0.0002	0.0007
carrier	0	0.0005



Studfarm / distribute evidence (8/8)

$$q_{ABDE \rightarrow BCE} := (q_{ADEI \rightarrow ABDE} \cdot p(D|A, B))^{\downarrow BE} = \begin{array}{c|cc} E & \text{pure carrier} \\ \hline B = \text{pure carrier} & 0.0003 & 0.0009 \\ \hline & 0.0005 & 0.0319 \end{array}$$



Studfarm / marginalize to target domains (1/5)

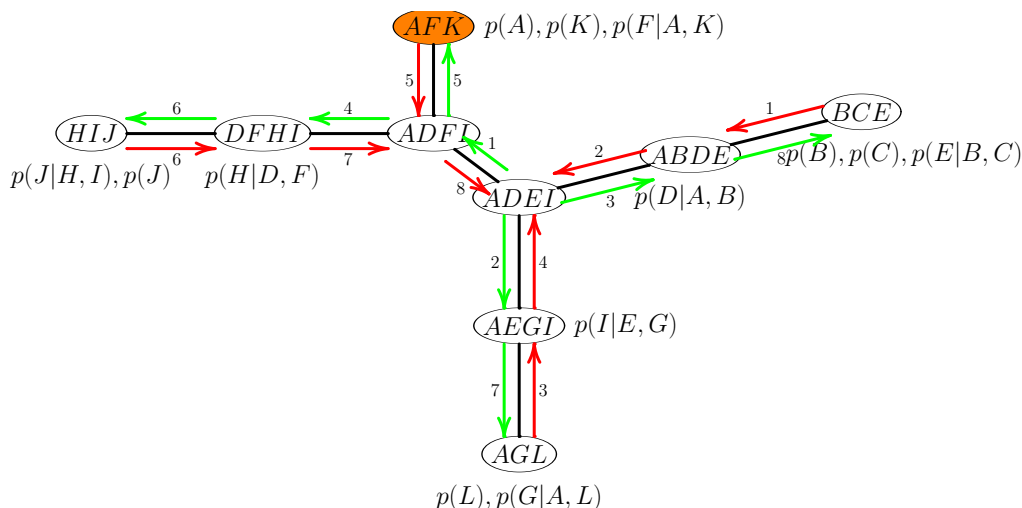
marginalize to target domains:

$$p_e(AFK) := q_{ADFI \rightarrow AFK} \cdot p(A) \cdot p(K) \cdot p(F|A, K)$$

$$p_e(A) := p_e(AFK)^{\downarrow A} = \begin{array}{c|cc} A = \text{pure carrier} & 0.3764 \\ \hline & 0.6236 \end{array}$$

$$p_e(F) := p_e(AFK)^{\downarrow F} = \begin{array}{c|cc} F = \text{pure carrier} & 0.5517 \\ \hline & 0.4483 \end{array}$$

$$p_e(K) := p_e(AFK)^{\downarrow K} = \begin{array}{c|cc} K = \text{pure carrier} & 0.9797 \\ \hline & 0.0203 \end{array}$$



Studfarm / marginalize to target domains (2/5)

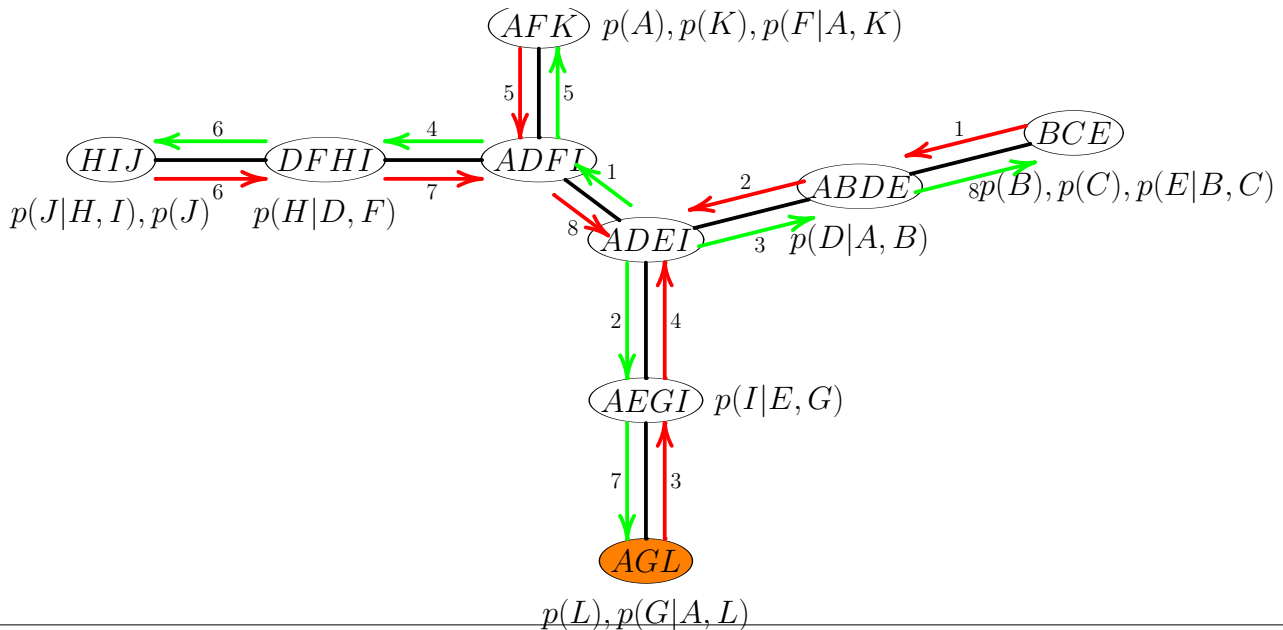
$$p_e(AGL) := q_{AEGI \rightarrow AGL} \cdot p(L) \cdot p(G|A, L)$$

$$p_e(G) := p_e(AGL) \downarrow^G =$$

$G = \text{pure carrier}$	0.375
	0.625

$$p_e(L) := p_e(AGL) \downarrow^L =$$

$L = \text{pure carrier}$	0.982
	0.018



Studfarm / marginalize to target domains (3/5)

$$p_e(BCE) := q_{ABDE \rightarrow BCE} \cdot p(B) \cdot p(C) \cdot p(E|B, C)$$

$$p_e(B) := p_e(BCE) \downarrow^B =$$

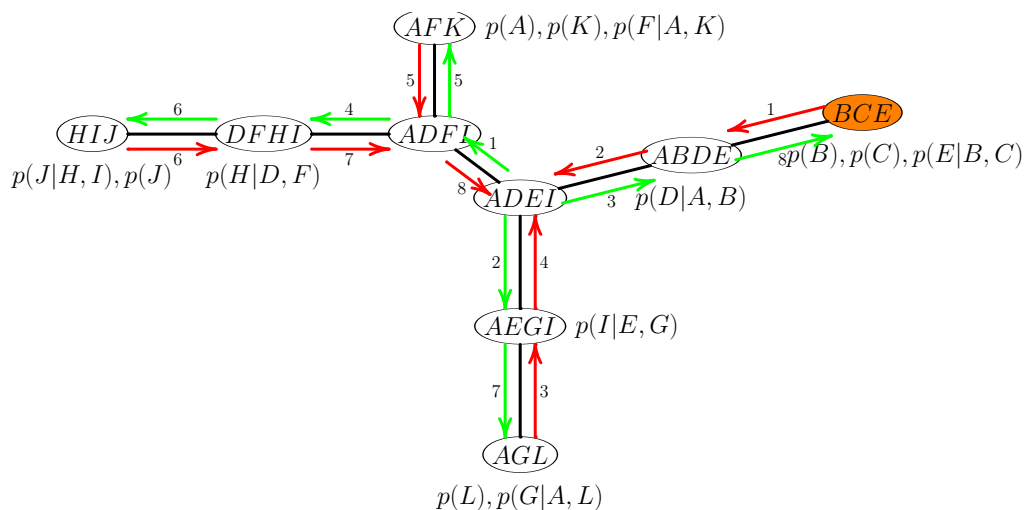
$B = \text{pure carrier}$	0.6192
	0.3808

$$p_e(C) := p_e(BCE) \downarrow^C =$$

$C = \text{pure carrier}$	0.9812
	0.0188

$$p_e(E) := p_e(BCE) \downarrow^E =$$

$E = \text{pure carrier}$	0.6138
	0.3862



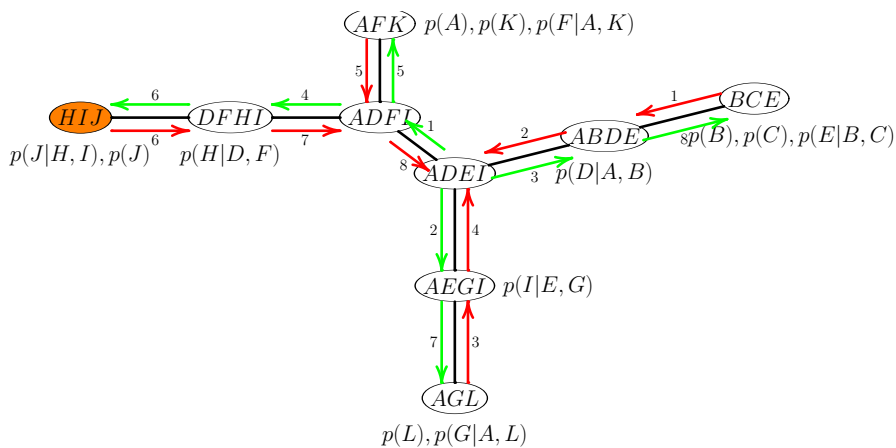
Studfarm / marginalize to target domains (4/5)

$$p_e(HIJ) := q_{DFHI \rightarrow HIJ} \cdot p(J|H, I) \cdot p(J)$$

$$p_e(H) := p_e(HIJ) \downarrow^H = \begin{array}{|l|l} H = \text{pure} & 0 \\ \text{carrier} & 1 \end{array}$$

$$p_e(I) := p_e(HIJ) \downarrow^I = \begin{array}{|l|l} I = \text{pure} & 0 \\ \text{carrier} & 1 \end{array}$$

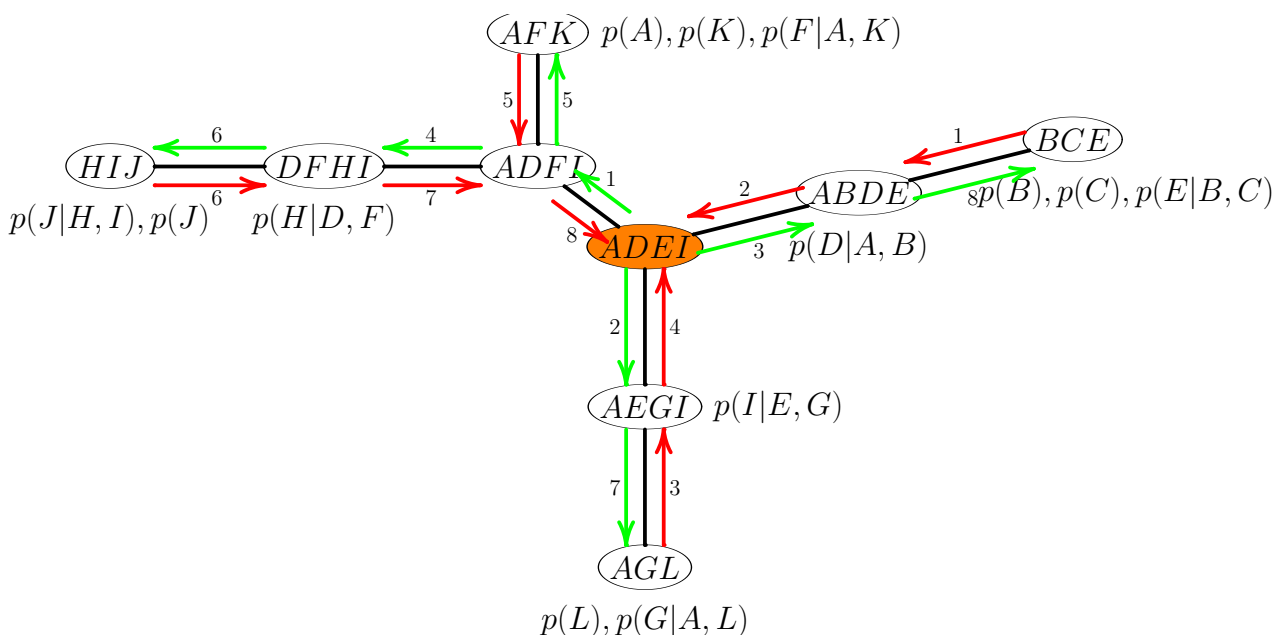
$$p_e(J) := p_e(HIJ) \downarrow^J = \begin{array}{|l|l} J = \text{pure} & 0 \\ \text{carrier} & 0 \\ \text{sick} & 1 \end{array}$$



Studfarm / marginalize to target domains (5/5)

$$p_e(ADEI) := q_{ABDE \rightarrow ADEI} \cdot q_{AEGI \rightarrow ADEI} \cdot q_{ADFI \rightarrow ADEI}$$

$$p_e(D) := p_e(ADEI) \downarrow^D = \begin{array}{|l|l} D = \text{pure} & 0.195 \\ \text{carrier} & 0.805 \end{array}$$



Studfarm / all single-variable marginals

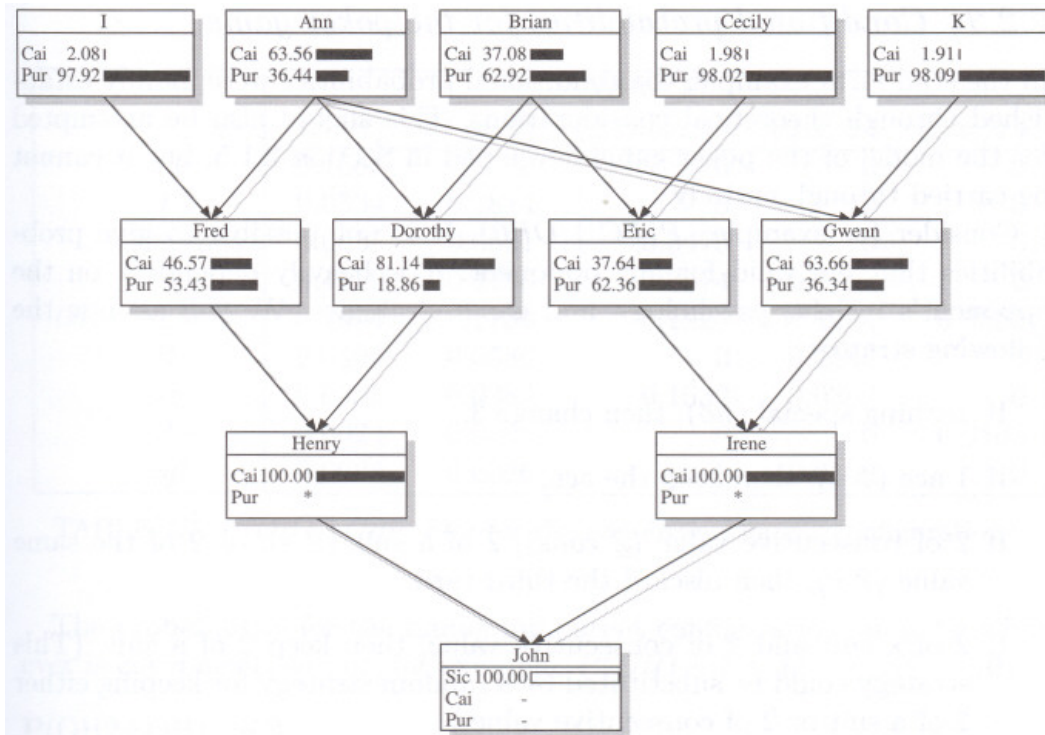


Figure 9: Probabilities given evidence that John is sick (AA). [Jen01, p. 49]

Studfarm / higher marginals

Say, we are interested not only in single-variable marginals, but in joint marginals of several variables, e.g., of A and B .

	B	pure	carrier
$p_e(A, B) =$	$A = \text{pure}$	0.00718	0.36919
	carrier	0.61205	0.01159

We proceed as follows:

If A and B are not conditionally independent given the evidence (here: J), then their joint potential is **not** the same as the product of $p_e(A)$ and $p_e(B)$, e.g.,

- (i) find a vertex potential that contains A and B ,
- (ii) compute that vertex potential from the link potentials,
- (iii) marginalize that vertex potential down to the target domain $\{A, B\}$.

$p_e(A) \cdot p_e(B) =$	$A = \text{pure}$	0.37636
	carrier	0.62364
	$B = \text{pure}$	0.61923
	carrier	0.38077

	B	pure	carrier
$=$	$A = \text{pure}$	0.23305	0.14331
	carrier	0.38617	0.23746

Hailfinder

Hailfinder has

- (i) 56 variables with
- (ii) 2 – 11 states: 2×2 , 25×3 , 20×4 , 2×5 , 3×6 , 2×7 and 2×11
- (iii) i.e., the hailfinder JPD has a total state space of size
119.307.129.289.316.404.700.753.638.195.200
- (iv) the clique tree found by MCS has a total state space size of 23050, that found by the minimal degree heuristics a total state space size of 9976.

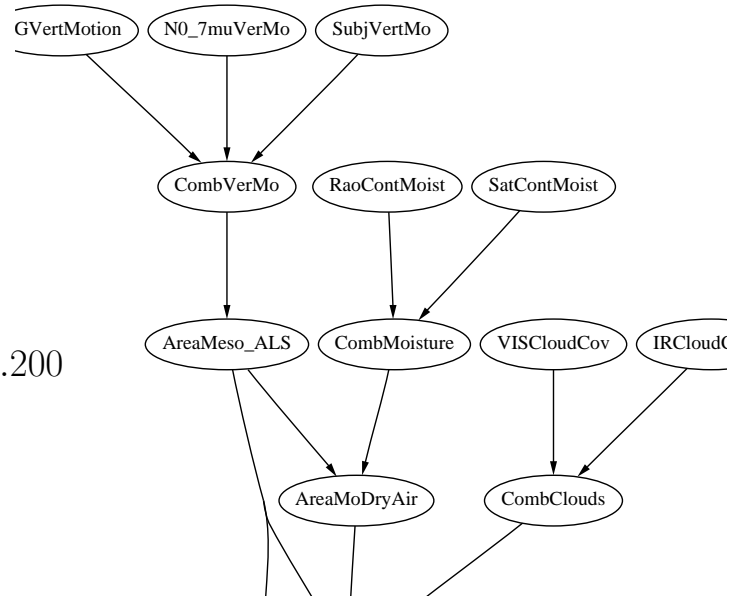


Figure 10: Hailfinder Bayesian network (detail view).

Hailfinder / clique tree (minimal degree)

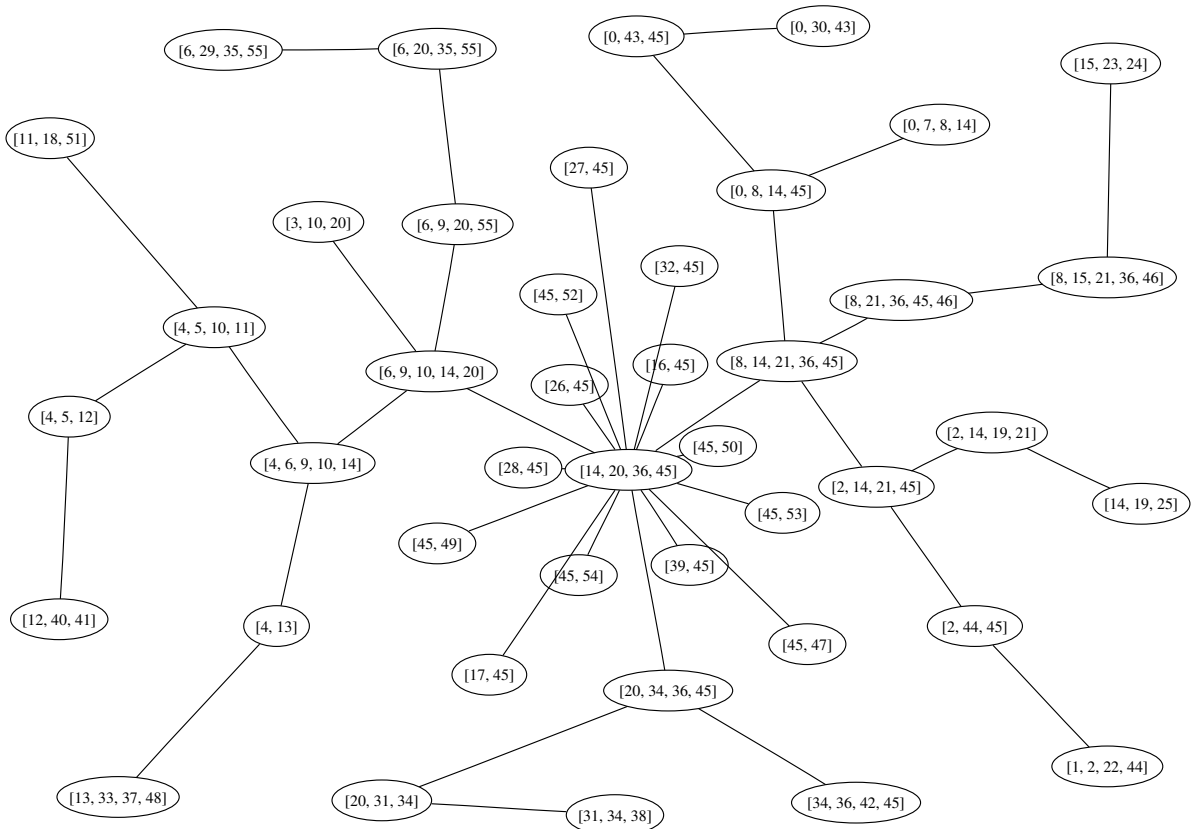


Figure 11: Clique tree of Hailfinder Bayesian network.

A full propagation of the Hailfinder network (including building the clique tree) takes

- (i) 4s for BNJ (Lauritzen-Spiegelhalter),
 - (ii) 1.2s for cgm-bn (Shafer-Shenoy; still not optimized)
- on a "standard notebook".

Hailfinder is small compared to many other published examples from the literature as

- (i) the Diabetes network with 413 variables,
- (ii) the Link network with 724 variables, or
- (iii) the Munin4 network with 1041 variables.

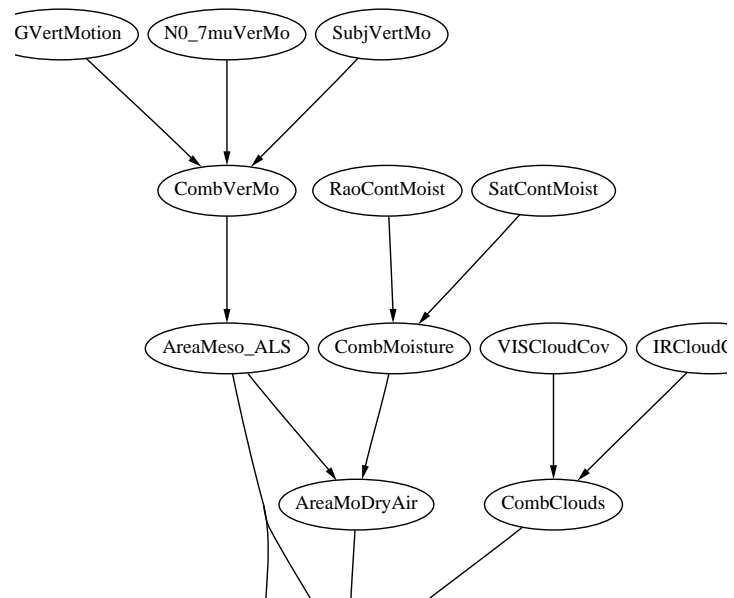


Figure 12: Hailfinder bayesian network (detail view).

References

- [Jen01] Finn V. Jensen. *Bayesian networks and decision graphs*. Springer, New York, 2001.