Bayesian Networks

III. Exact Inference (section 4)

Prof. Dr. Lars Schmidt-Thieme, L. B. Marinho, K. Buza

Information Systems and Machine Learning Lab(ISMLL)
University of Hildesheim

Studfarm / problem

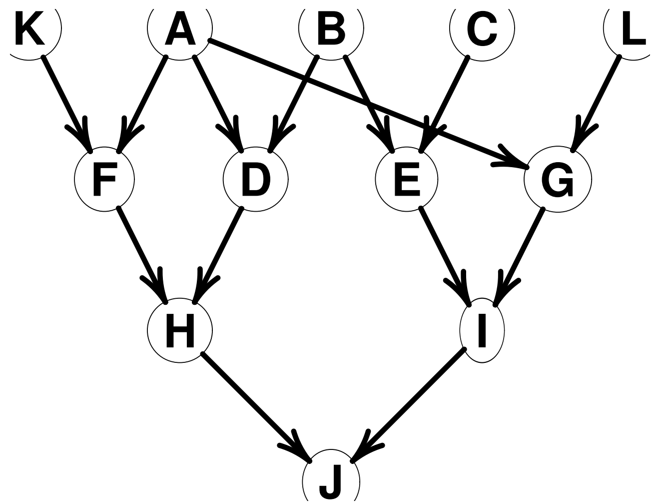


Figure 1: Studfarm bayesian network.

potentials:

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

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

(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	$\frac{1}{3}$
aA	0	.5	.5	$\frac{2}{3}$

(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 $p(J)$:

$$p(J = \text{AA}) = 1$$

Studfarm / markov network

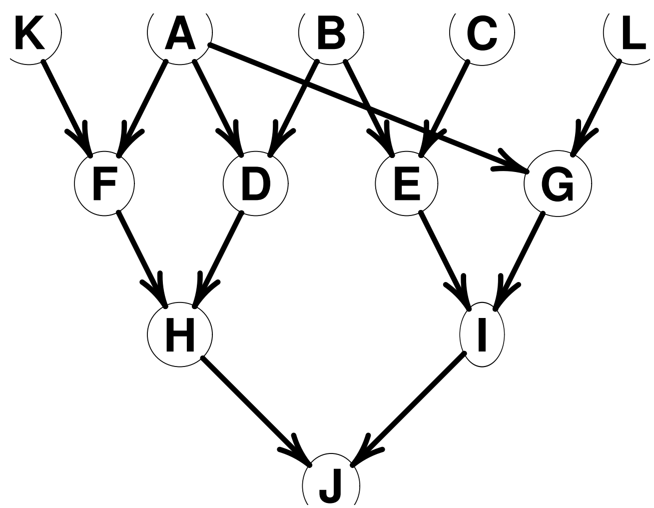


Figure 2: Studfarm bayesian network.

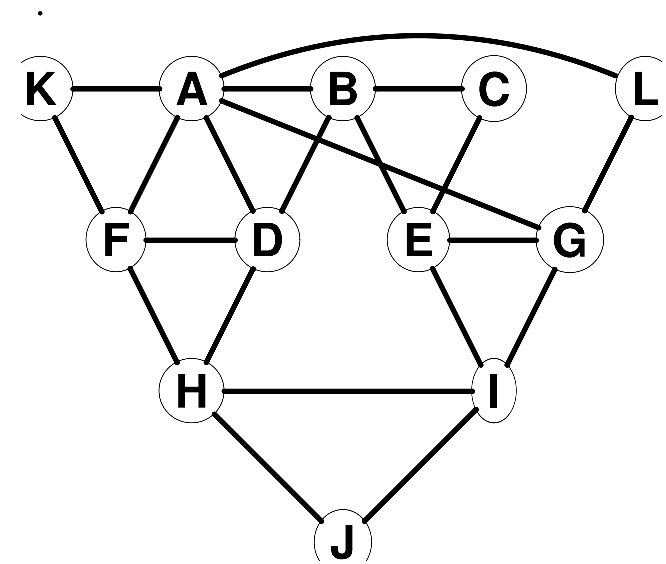


Figure 3: Studfarm markov network (moral graph).

Studfarm / triangulation & cluster tree (MCS)

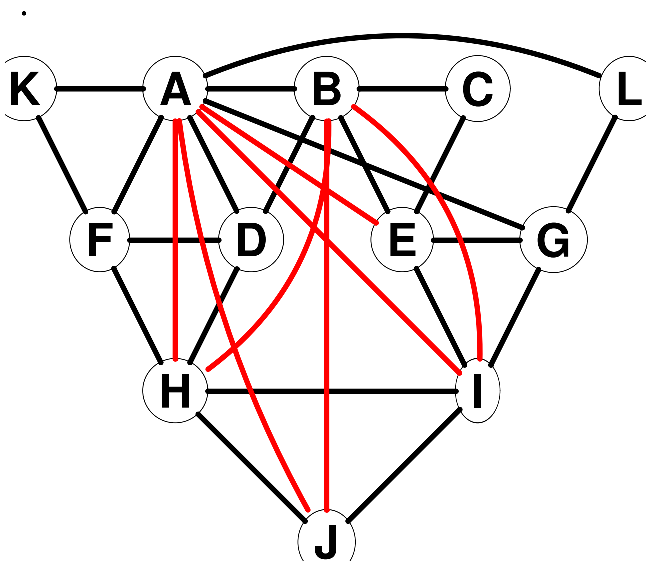


Figure 4: Triangulation of Studfarm markov network by MCS (fill-in 7).

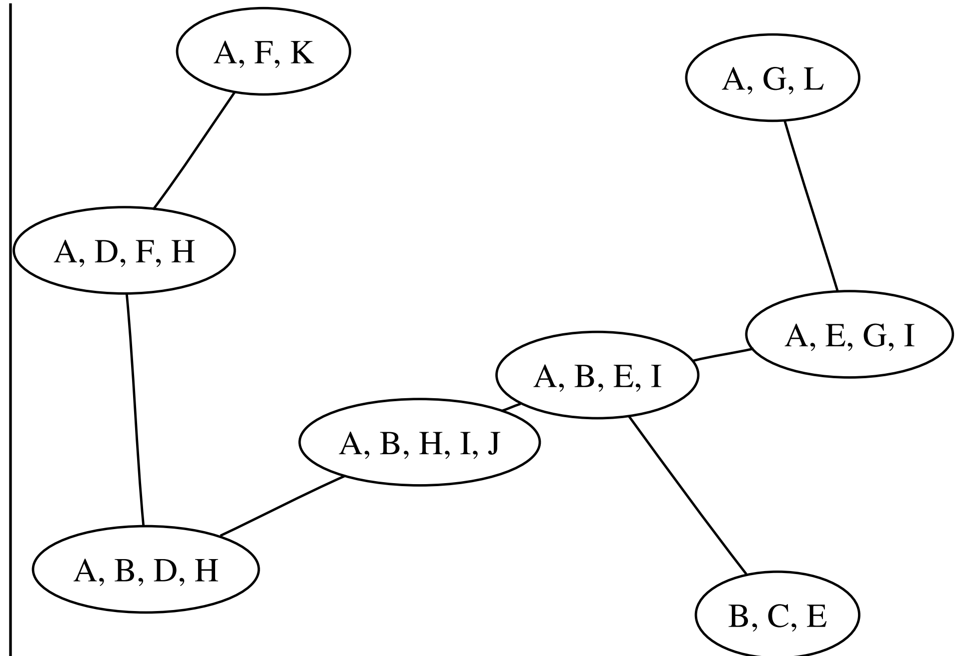


Figure 5: Cluster tree for the triangulation at the left (total state space size 136).

Studfarm / triangulation & cluster tree (Minimal degree)

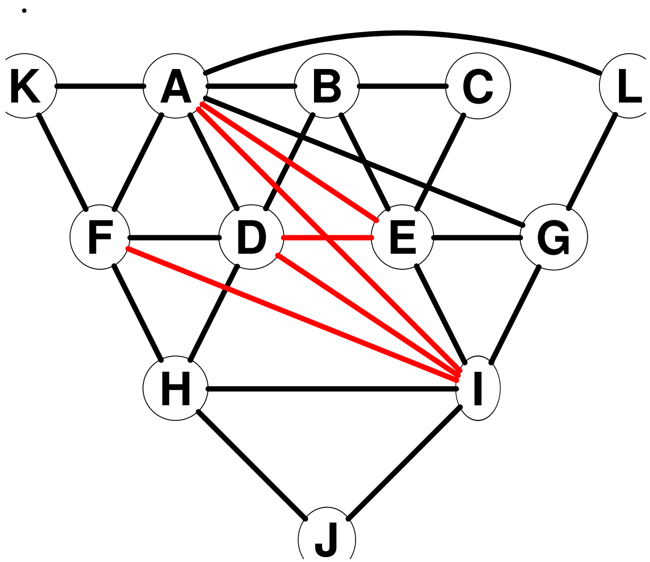


Figure 6: Triangulation of Studfarm markov network by Minimal Degree Heuristics (fill-in 5).

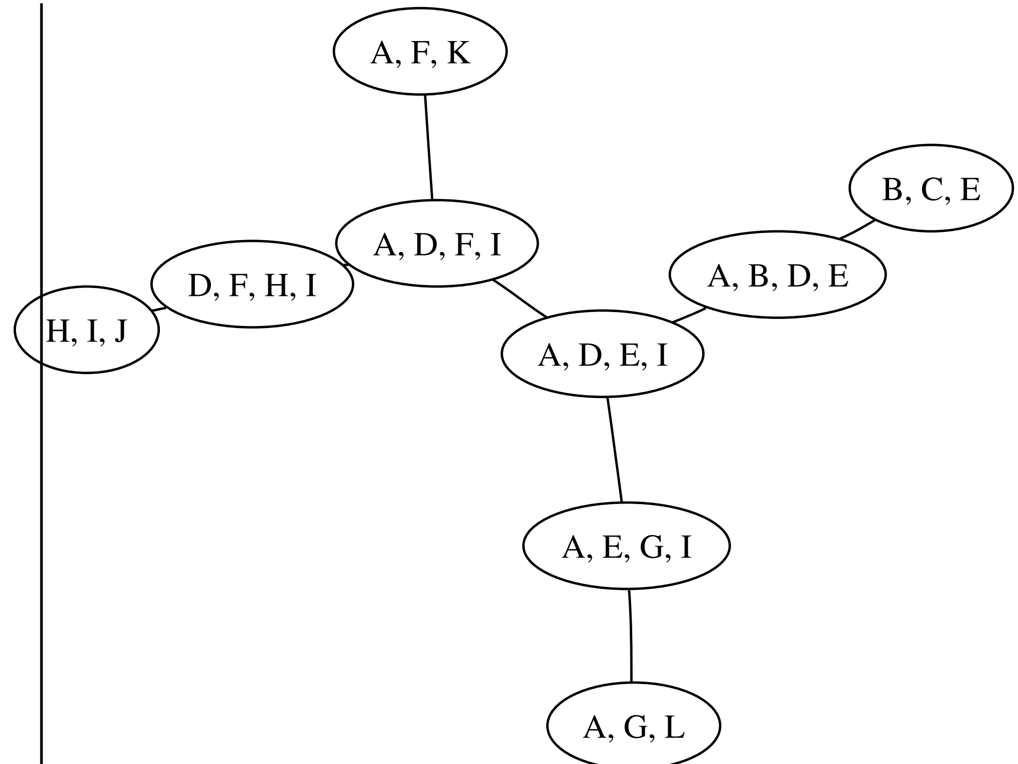


Figure 7: Cluster tree for the triangulation at the left (total state space size 116).

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 $p(J)$.

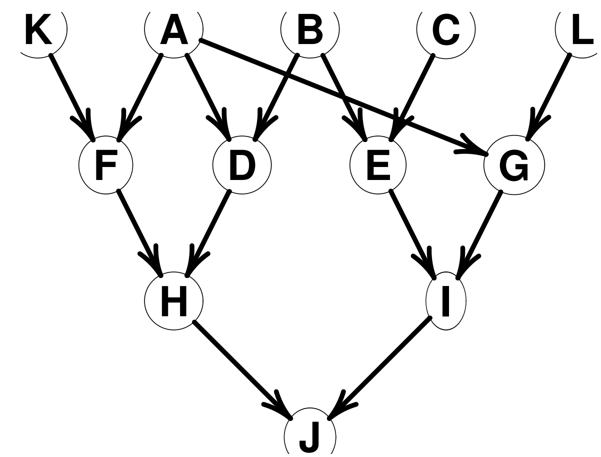
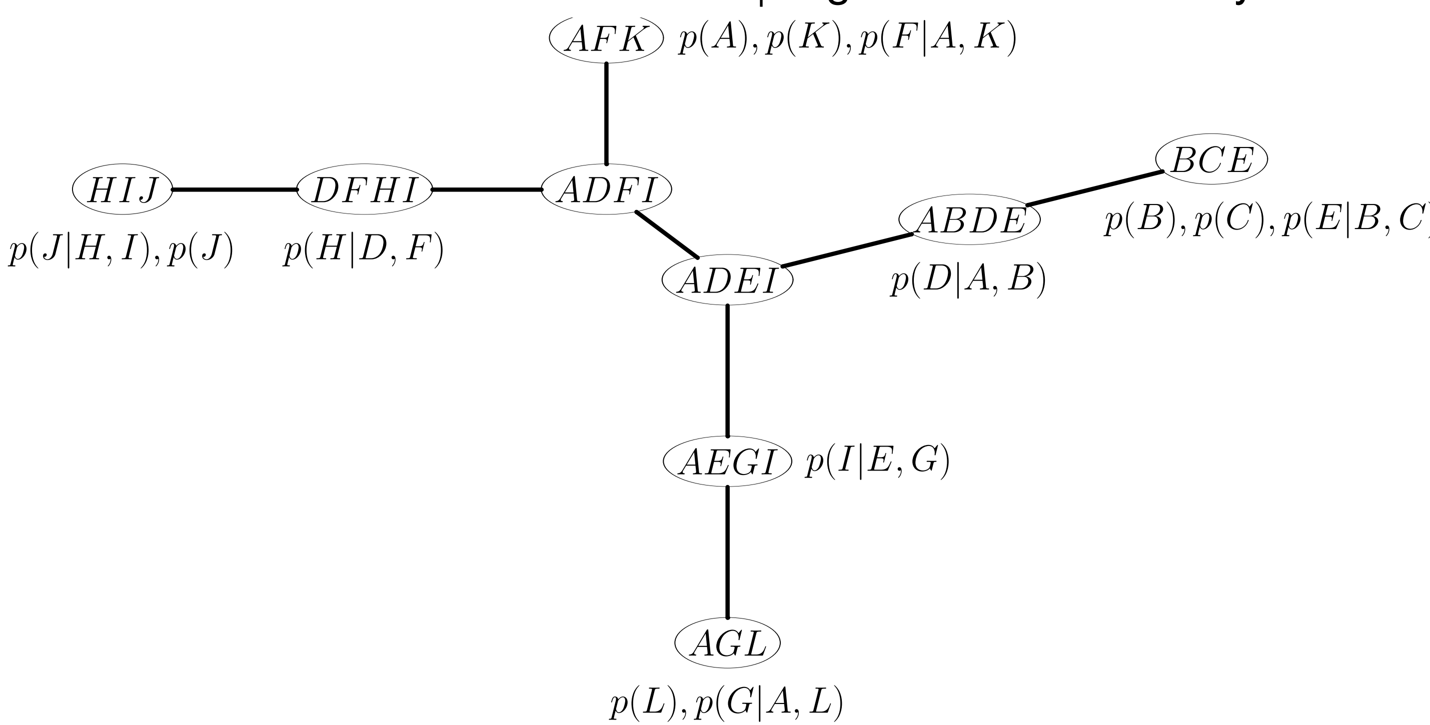


Figure 8: Studfarm bayesian network.

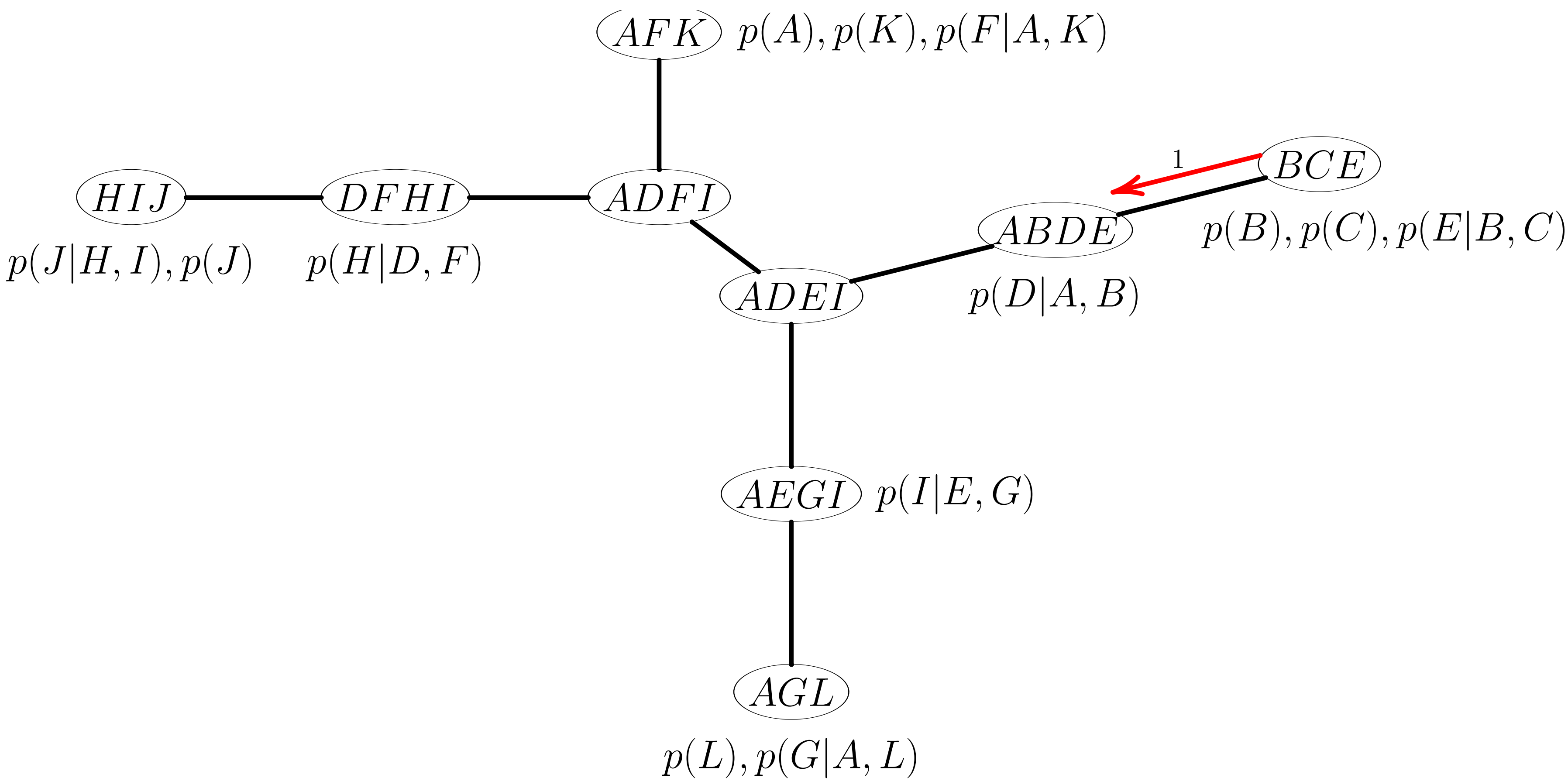


Studfarm / collect evidence (1/8)

collect evidence:

$$q_{BCE \to ABDE} := (p(B) \cdot p(C) \cdot p(E|B,C))^{\downarrow BE} =$$

E	pure	carrier
B = pure	0.985	0.005
carrier	0.005	0.005

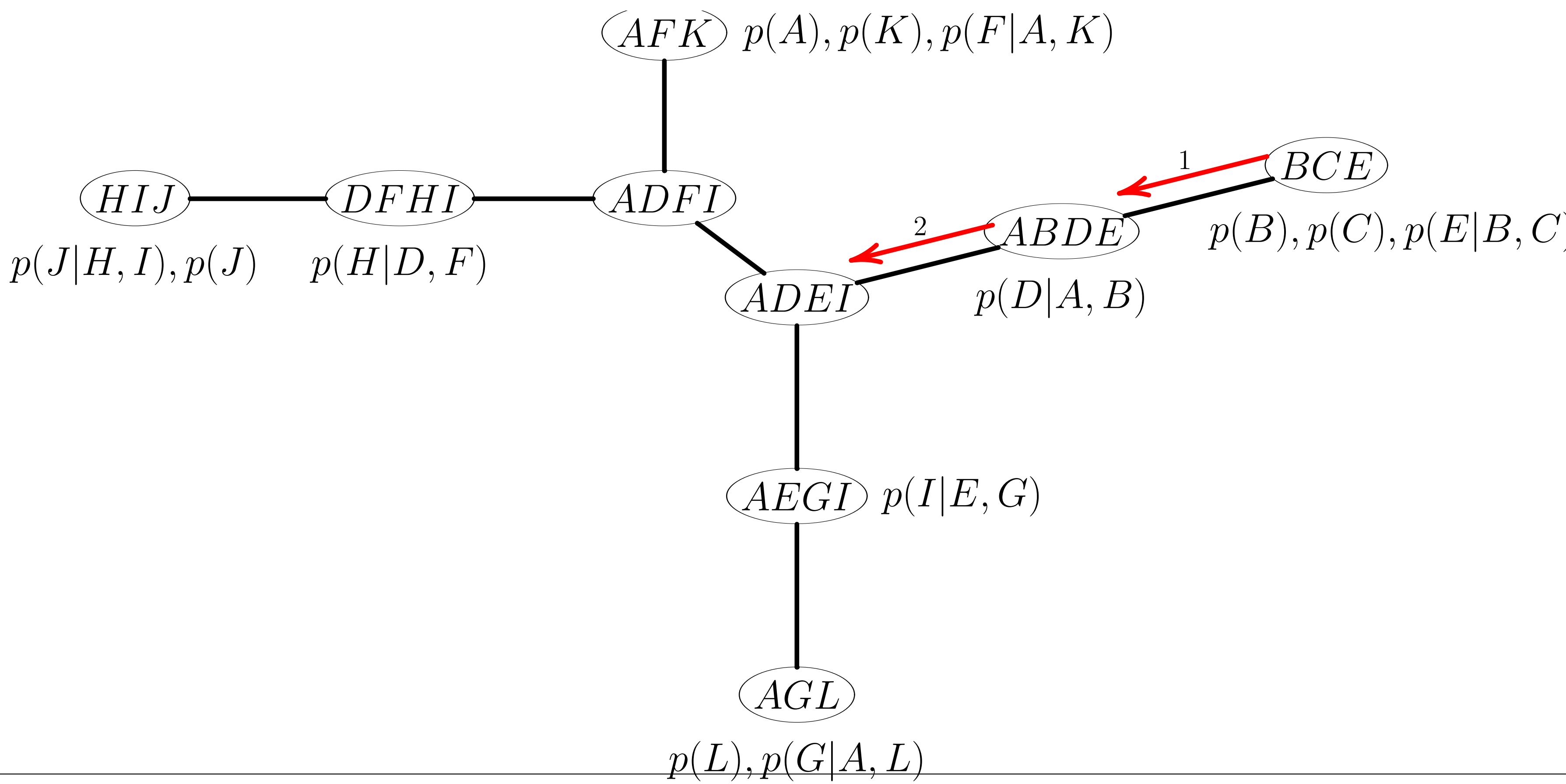


Studfarm / collect evidence (2/8)

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

=

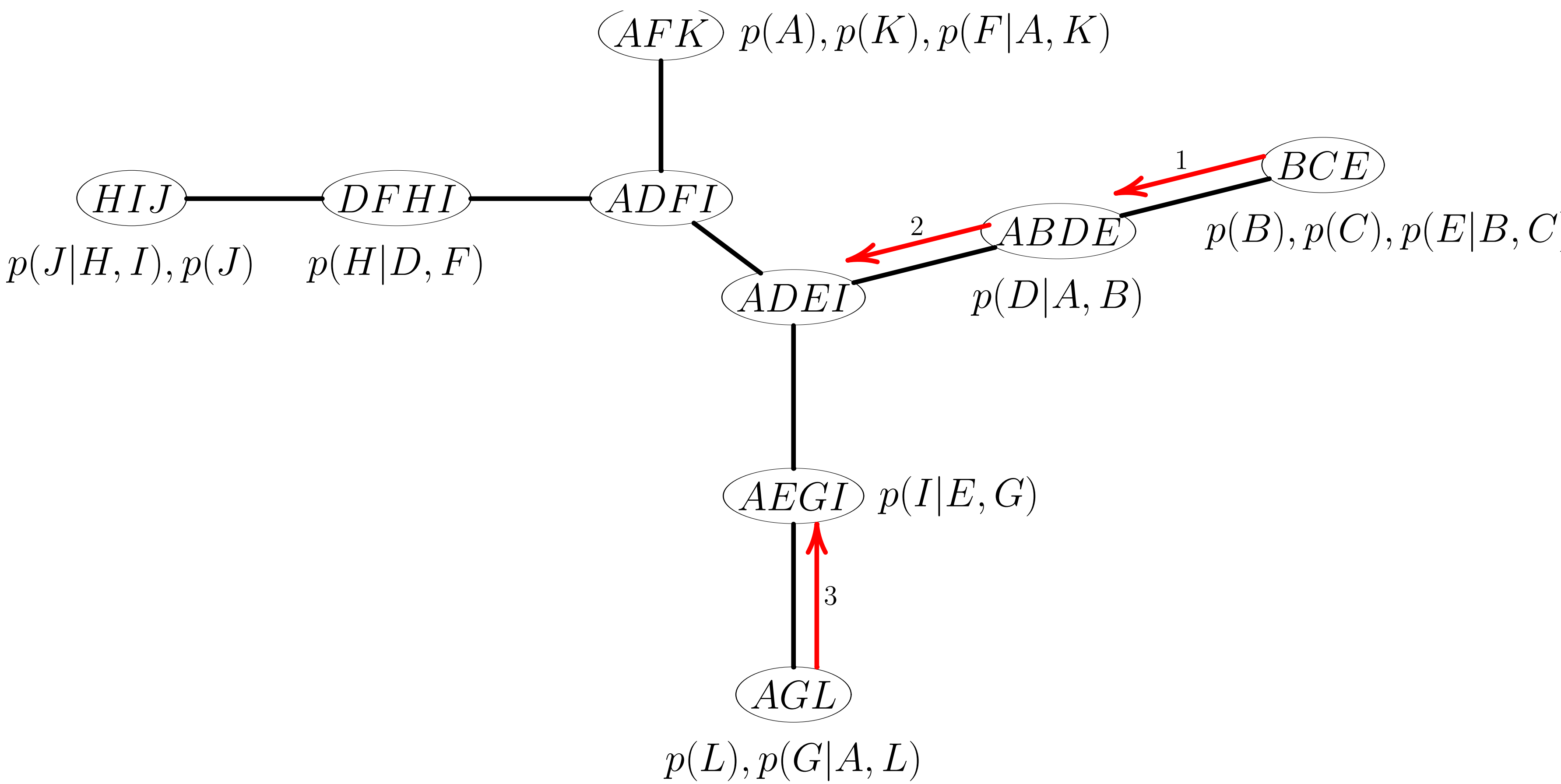
D	pure		carrier	
E	pure	carrier	pure	carrier
$A =$ pure	0.9875	0.0075	0.0025	0.0025
carrier	0.4942	0.0041	0.4958	0.0058



Studfarm / collect evidence (3/8)

$$q_{AGL \to AEGI} := (p(C) \cdot p(G|A, L))^{\downarrow AG} =$$

G	pure	carrier
$A =$ pure	0.995	0.005
carrier	0.4983	0.5017

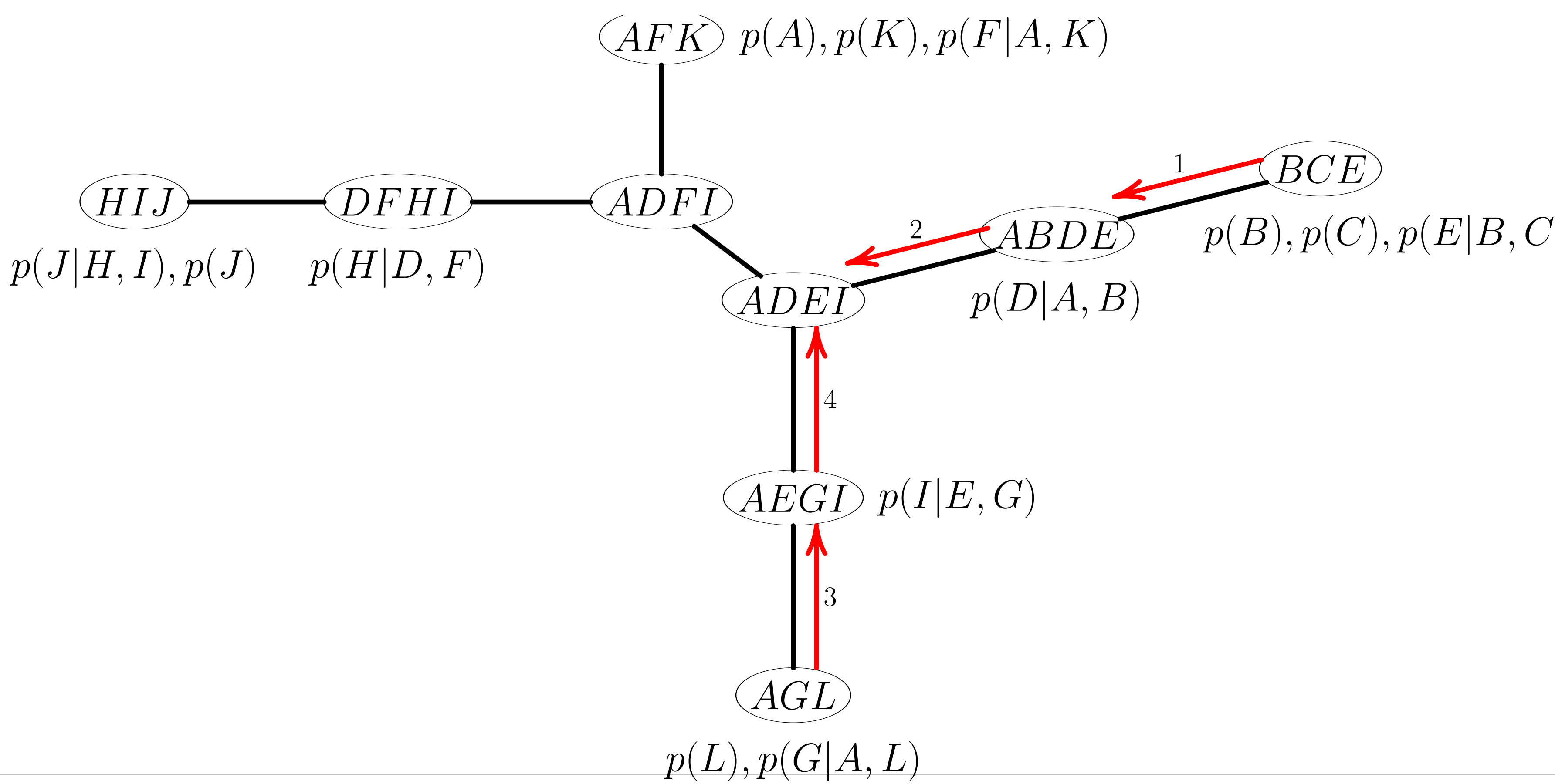


Studfarm / collect evidence (4/8)

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

=

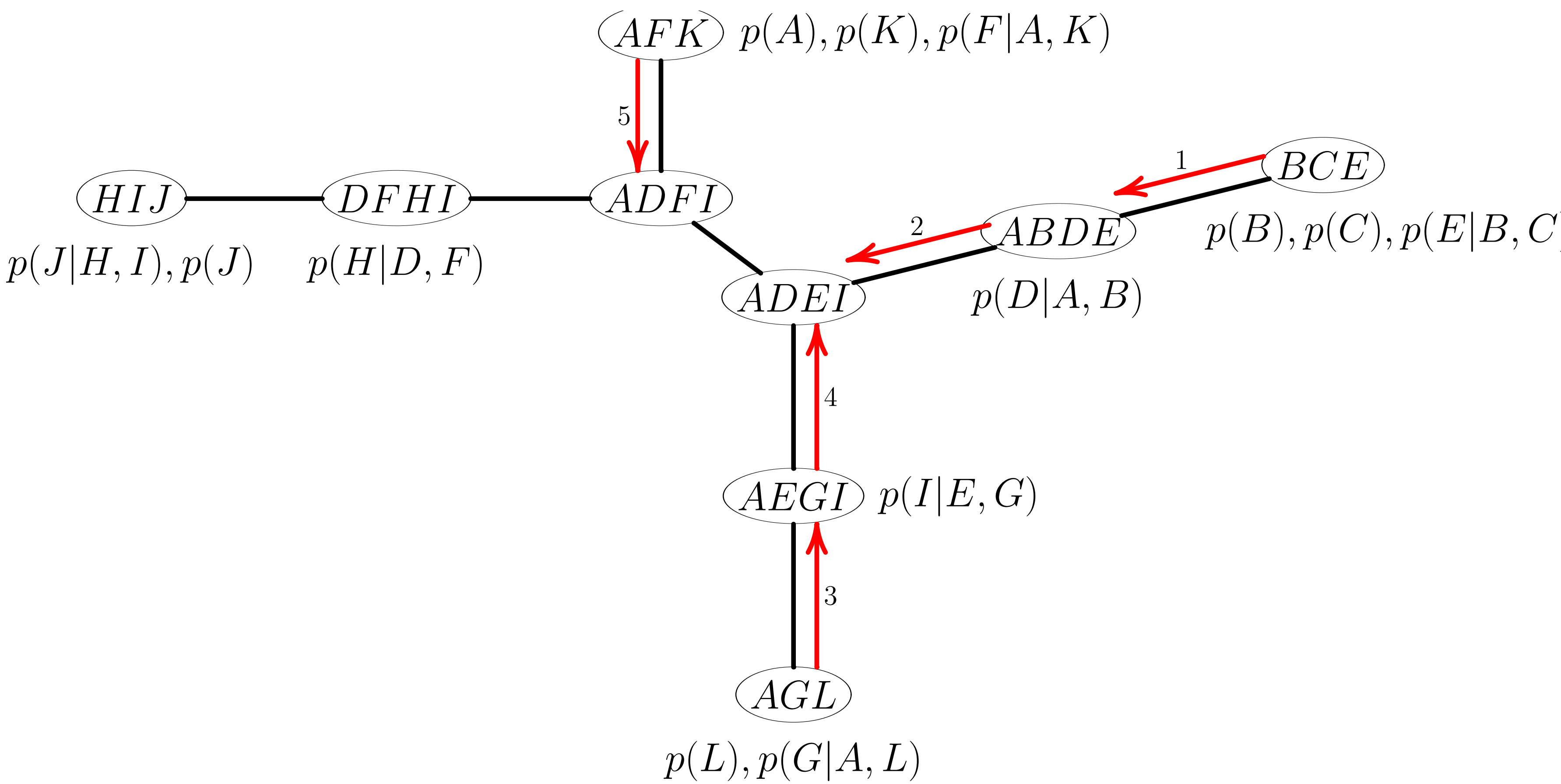
E	pure		carrier	
I	pure	carrier		
$A =$ pure	0.9975	0.0025	0.4992	0.5008
carrier	0.7492	0.2508	0.4164	0.5836



Studfarm / collect evidence (5/8)

$$q_{AFK \to ADFI} := (p(A) \cdot p(K) \cdot p(F|A,K))^{\downarrow AF} =$$

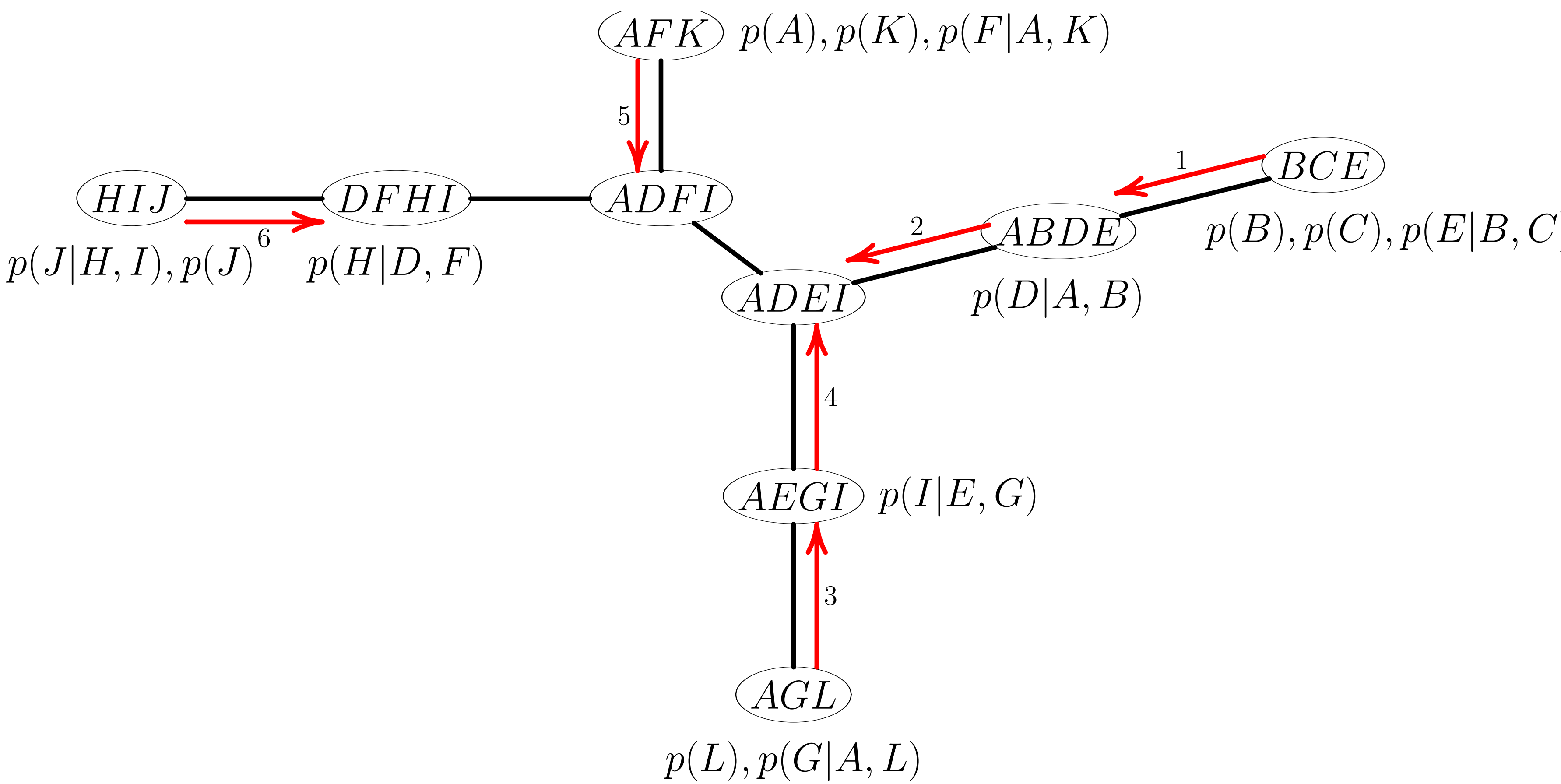
F	pure	carrier
$A =$ pure	0.985	0.005
carrier	0.005	0.005



Studfarm / collect evidence (6/8)

$$q_{HIJ \to DFHI} := (p(J|H,I) \cdot p(J))^{\downarrow HI} =$$

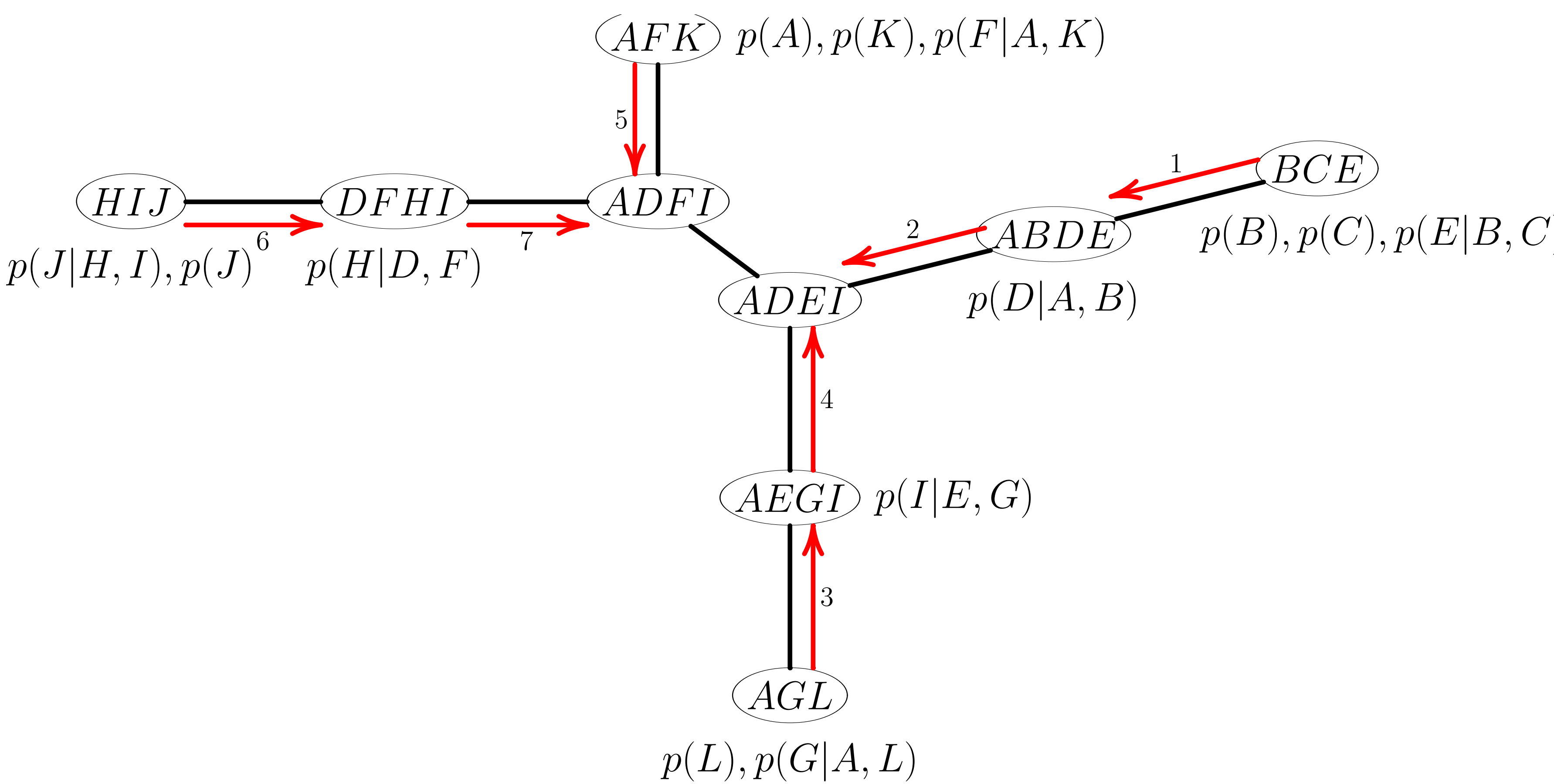
	I	pure	carrier
$H =$	pure	0	0
	carrier	0	0.25



Studfarm / collect evidence (7/8)

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

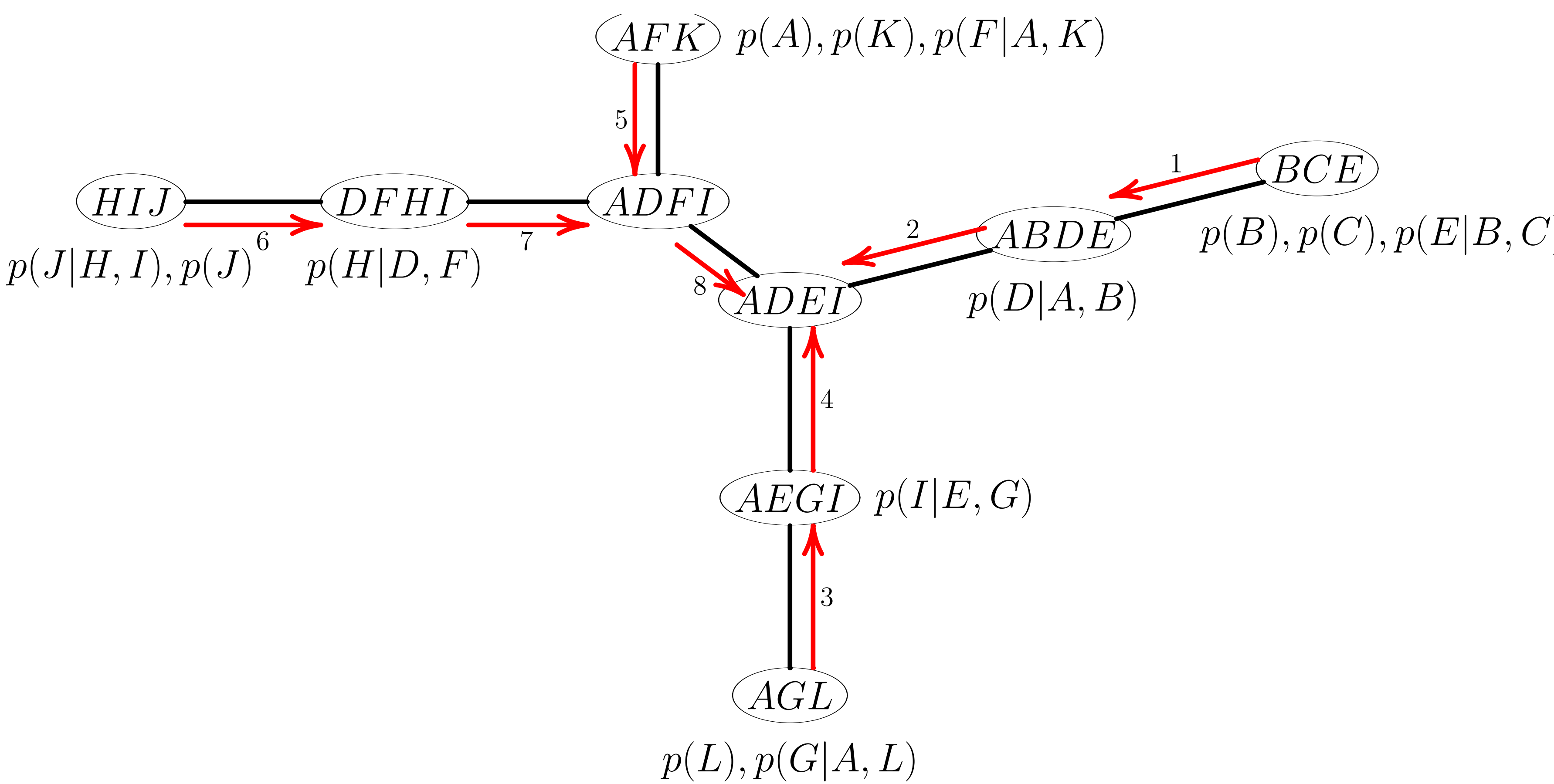
	F	pure		carrier	
	I	pure	carrier	pure	carrier
$D =$	pure	0	0	0	0.125
	carrier	0	0.125	0	0.1667



Studfarm / collect evidence (8/8)

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

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



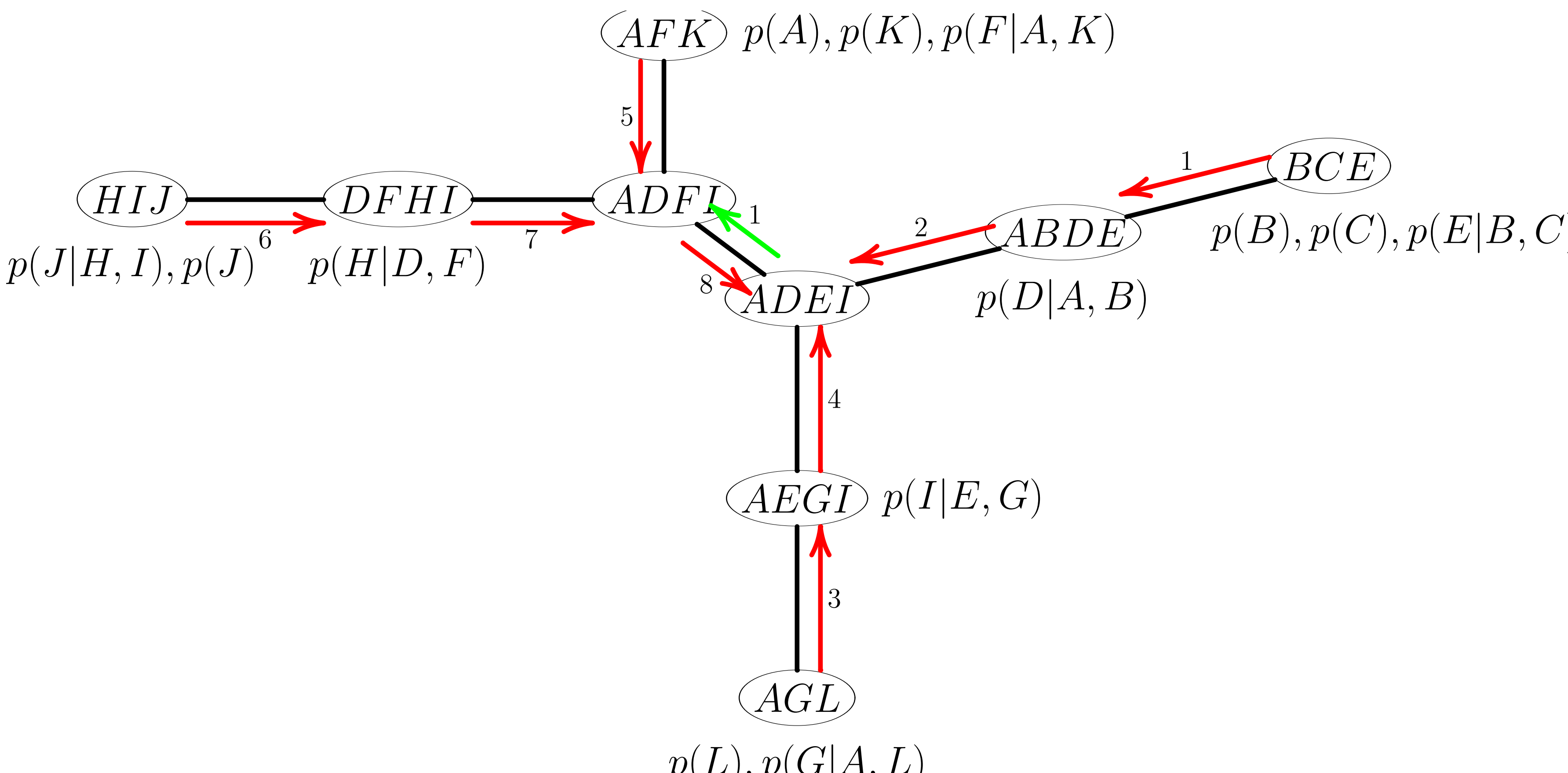
Studfarm / distribute evidence (1/8)

distribute evidence:

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

$=$

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

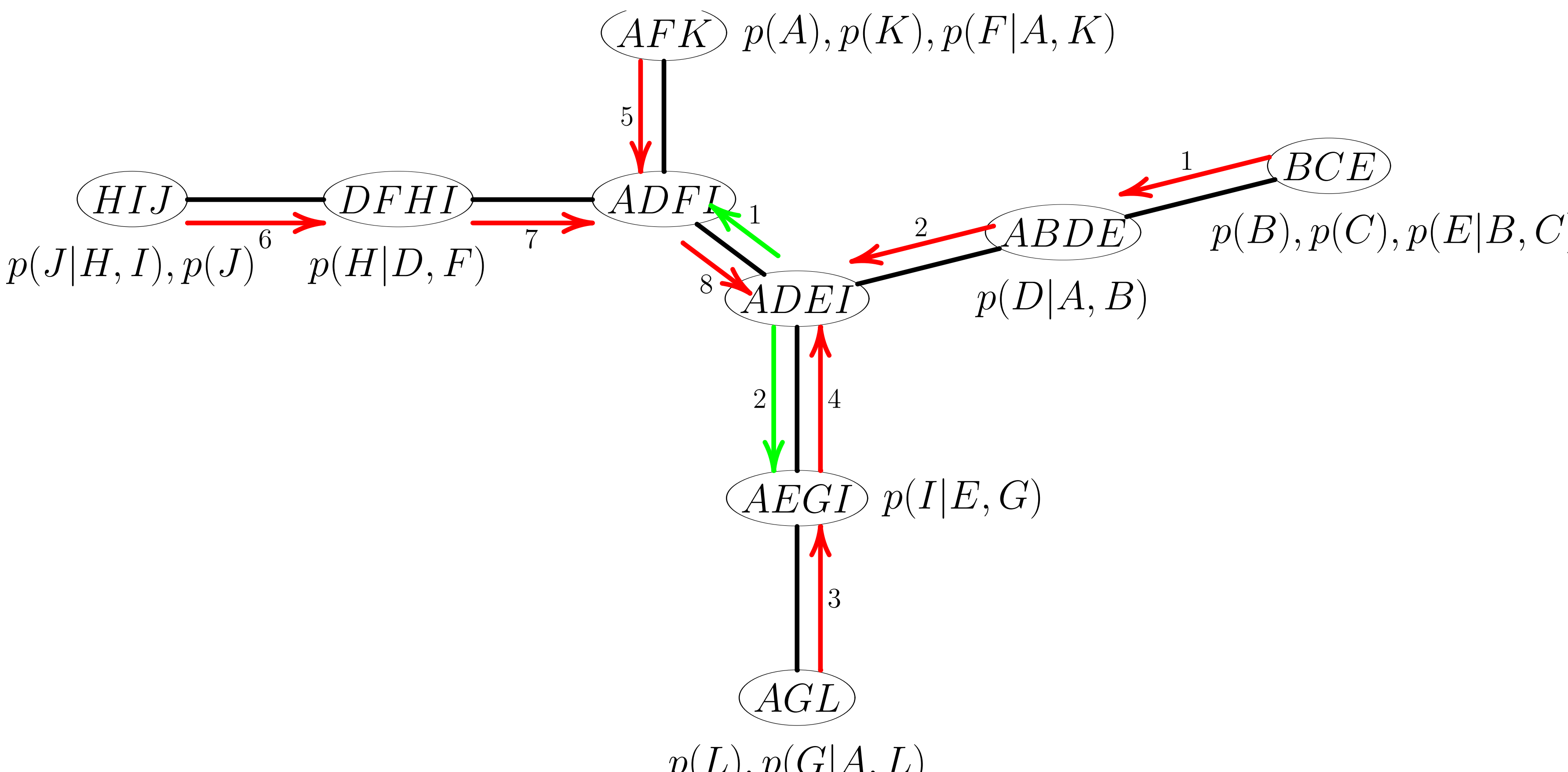


Studfarm / distribute evidence (2/8)

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

=

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

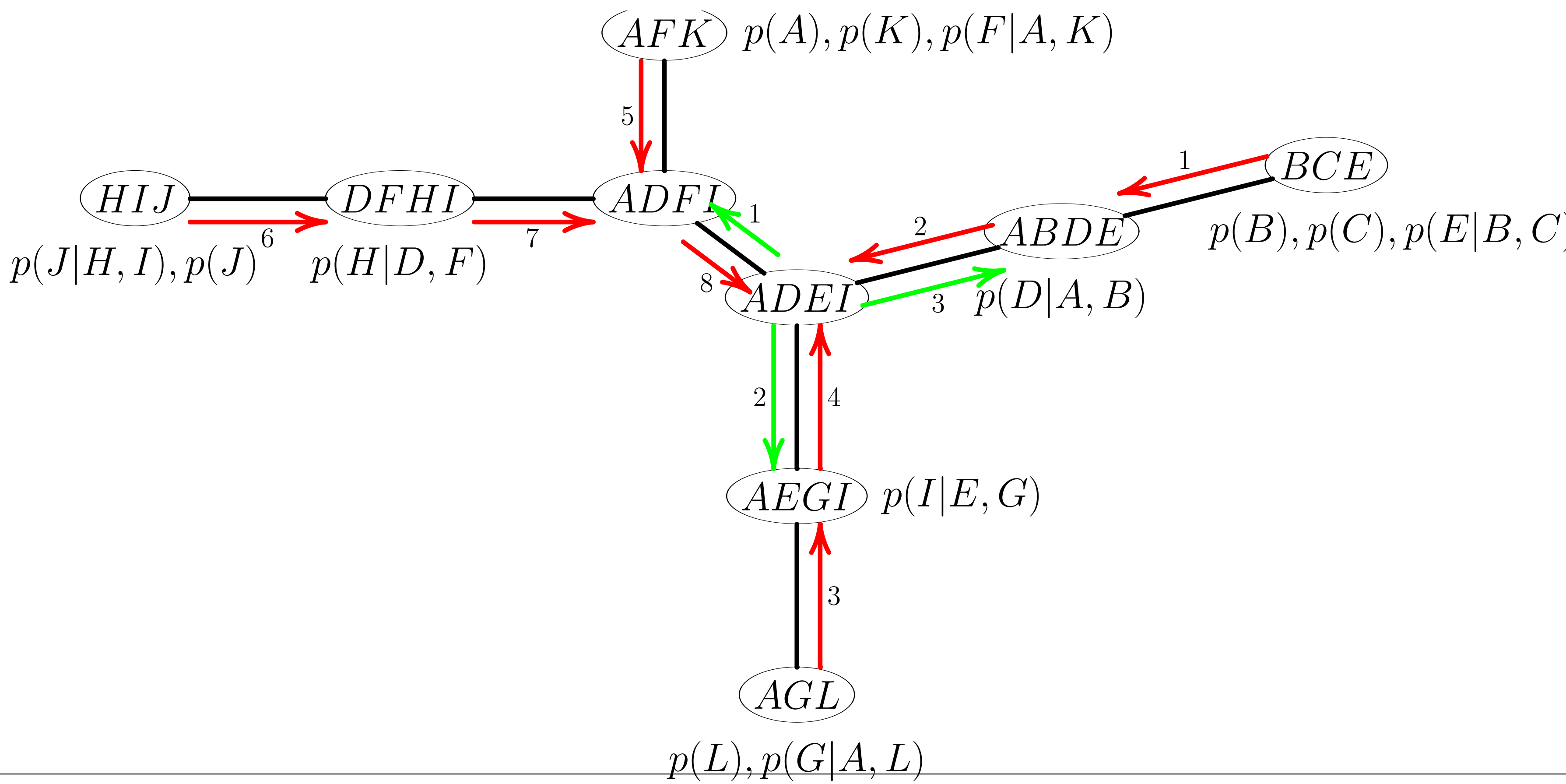


Studfarm / distribute evidence (3/8)

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

=

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

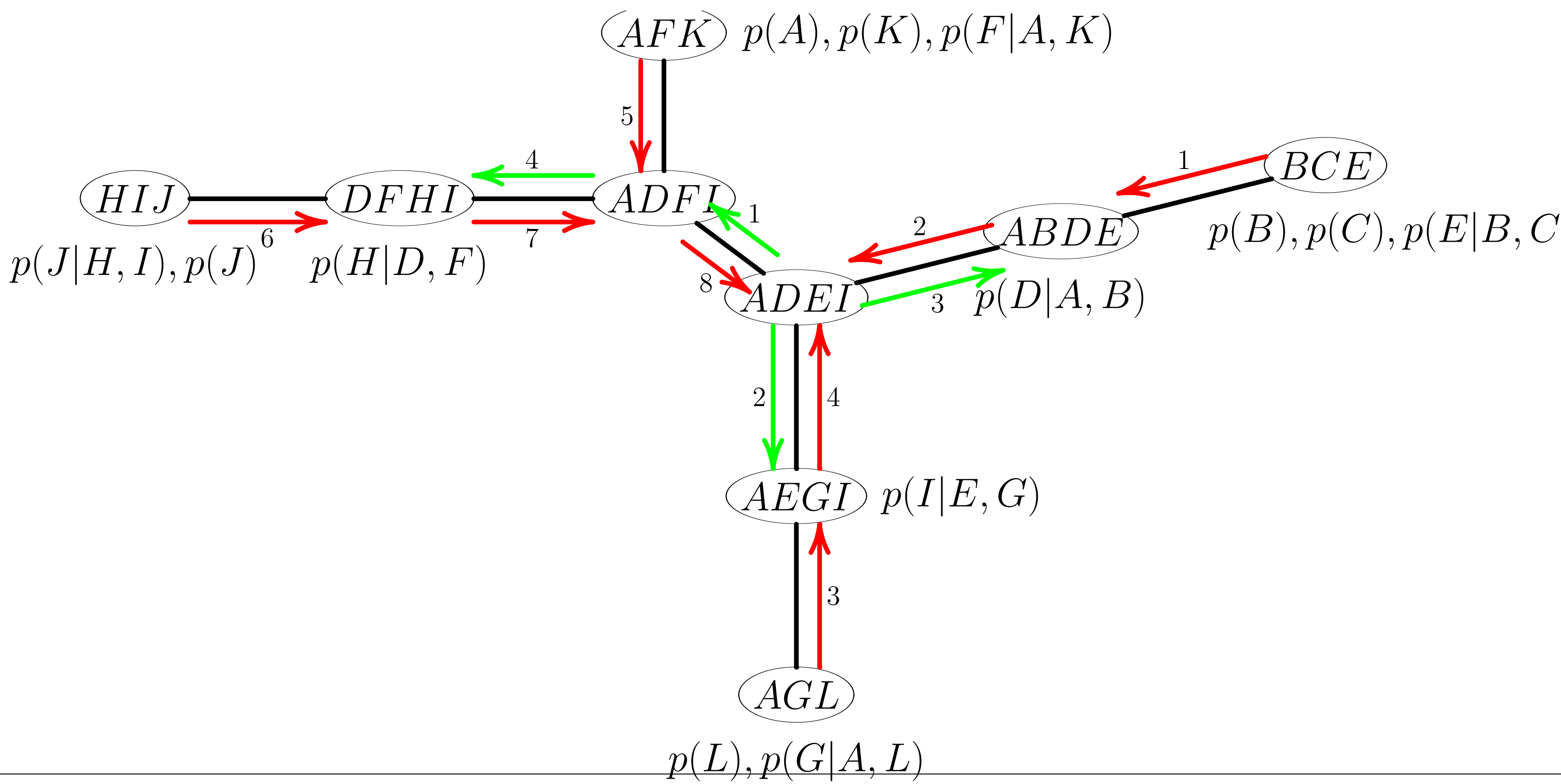


Studfarm / distribute evidence (4/8)

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

$=$

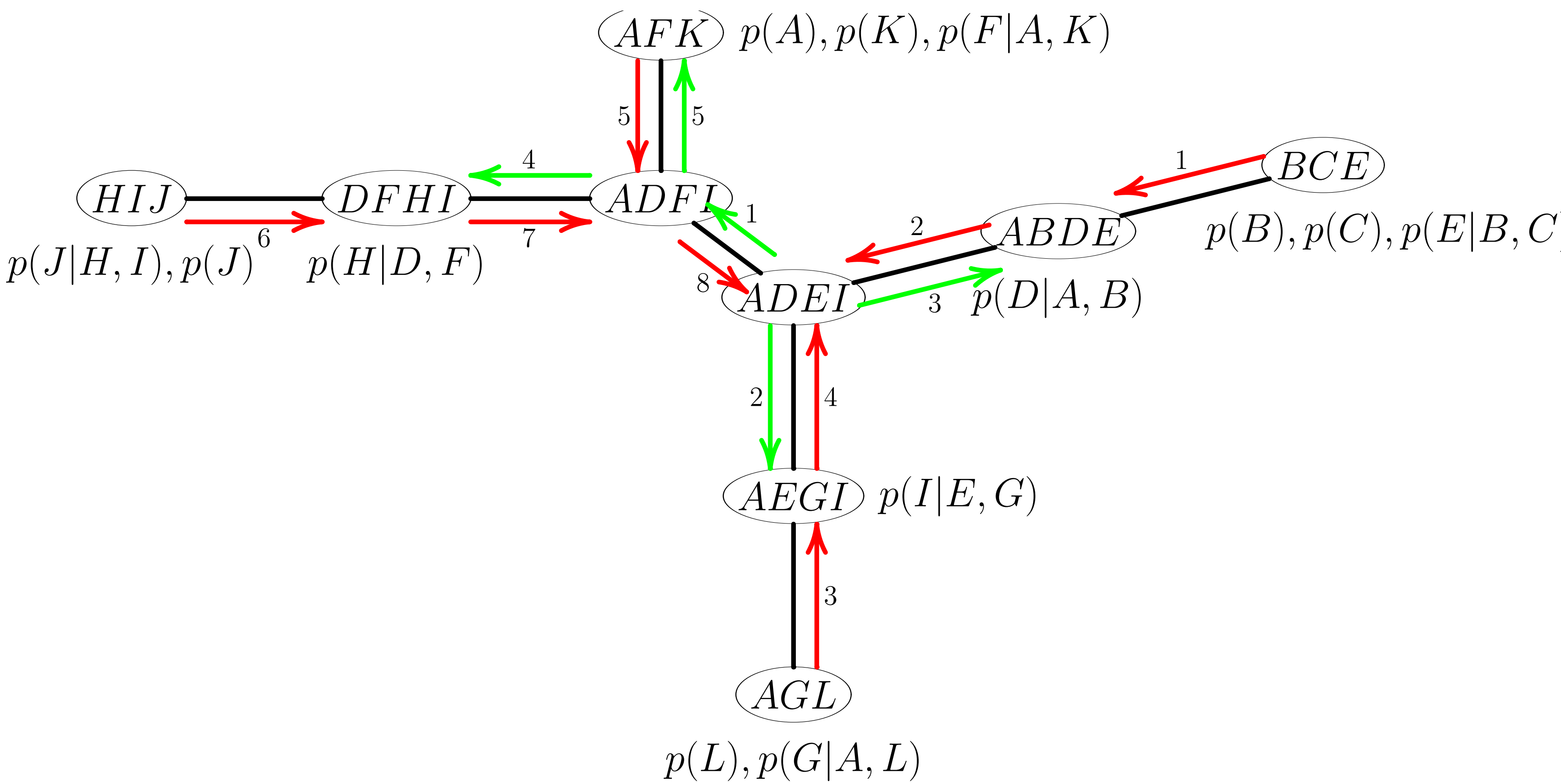
F	pure		carrier	
I	pure	carrier		
$D =$ pure	0.9759	0.0067	0.0068	0.0007
carrier	0.0055	0.0019	0.0019	0.0006



Studfarm / distribute evidence (5/8)

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

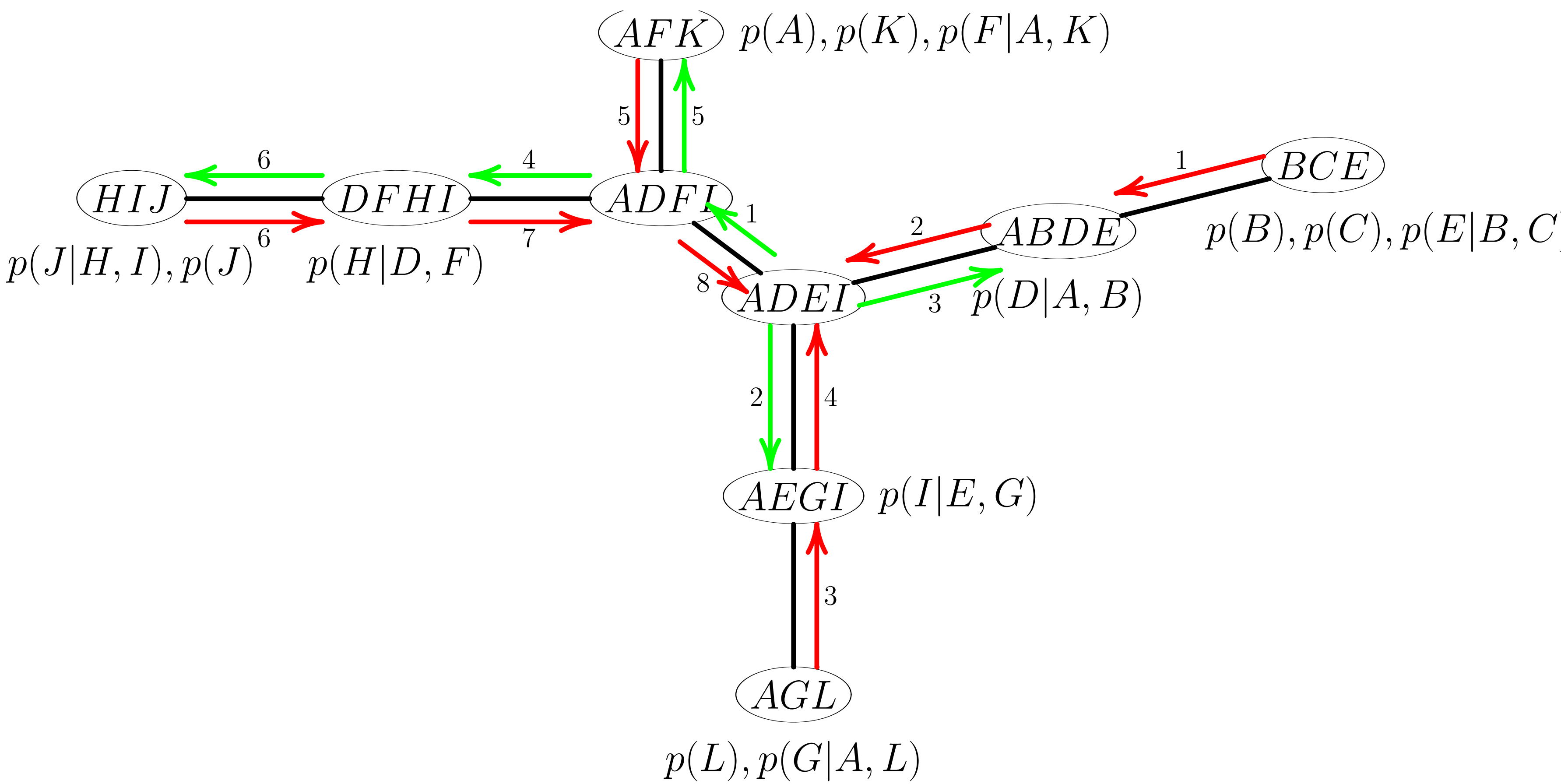
	F	pure	carrier
$A =$	pure	0.0002	0.001
	carrier	0.016	0.0371



Studfarm / distribute evidence (6/8)

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

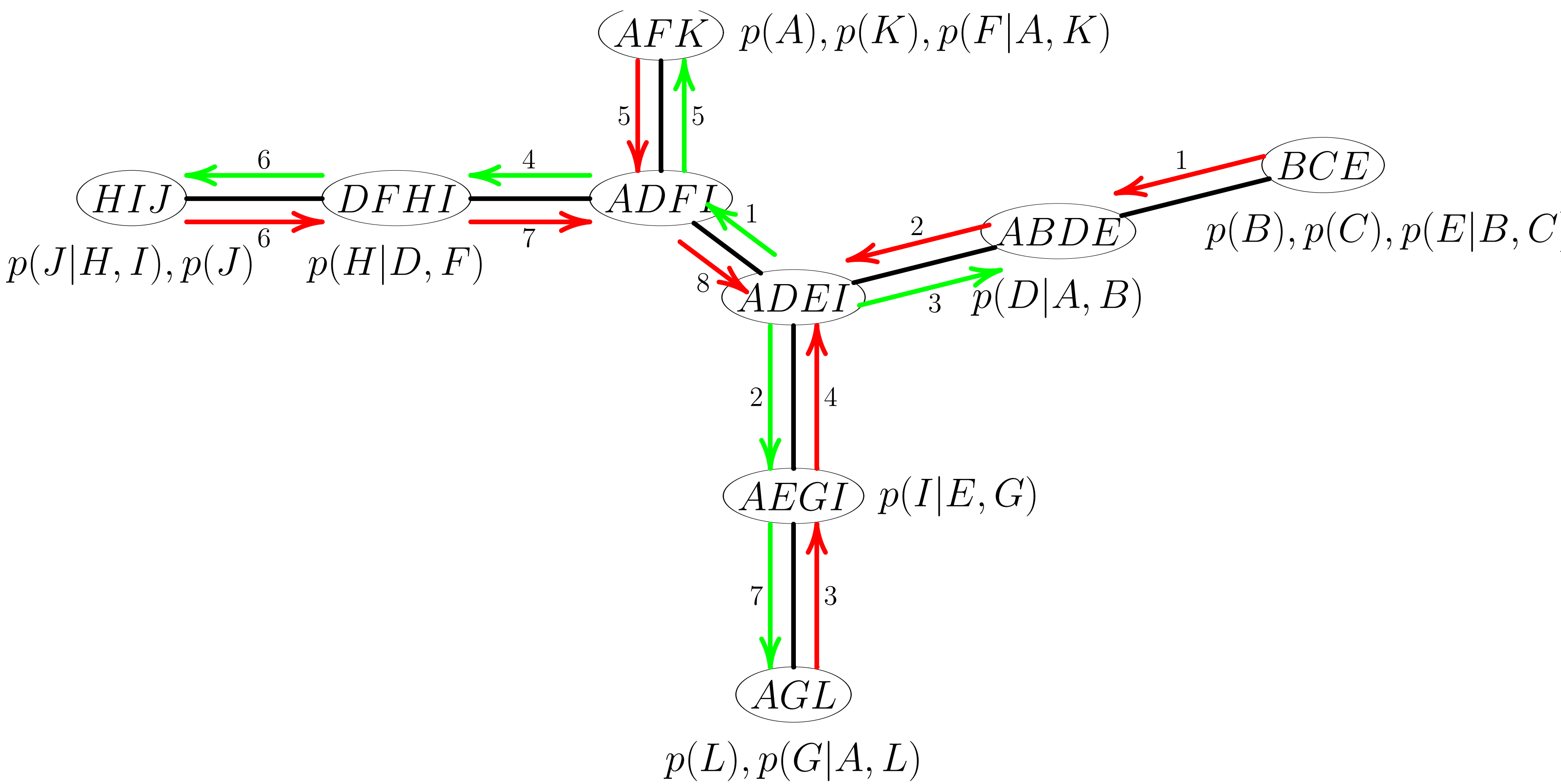
I	pure	carrier
$H =$ pure	0.9827	0.0082
carrier	0.0074	0.0017



Studfarm / distribute evidence (7/8)

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

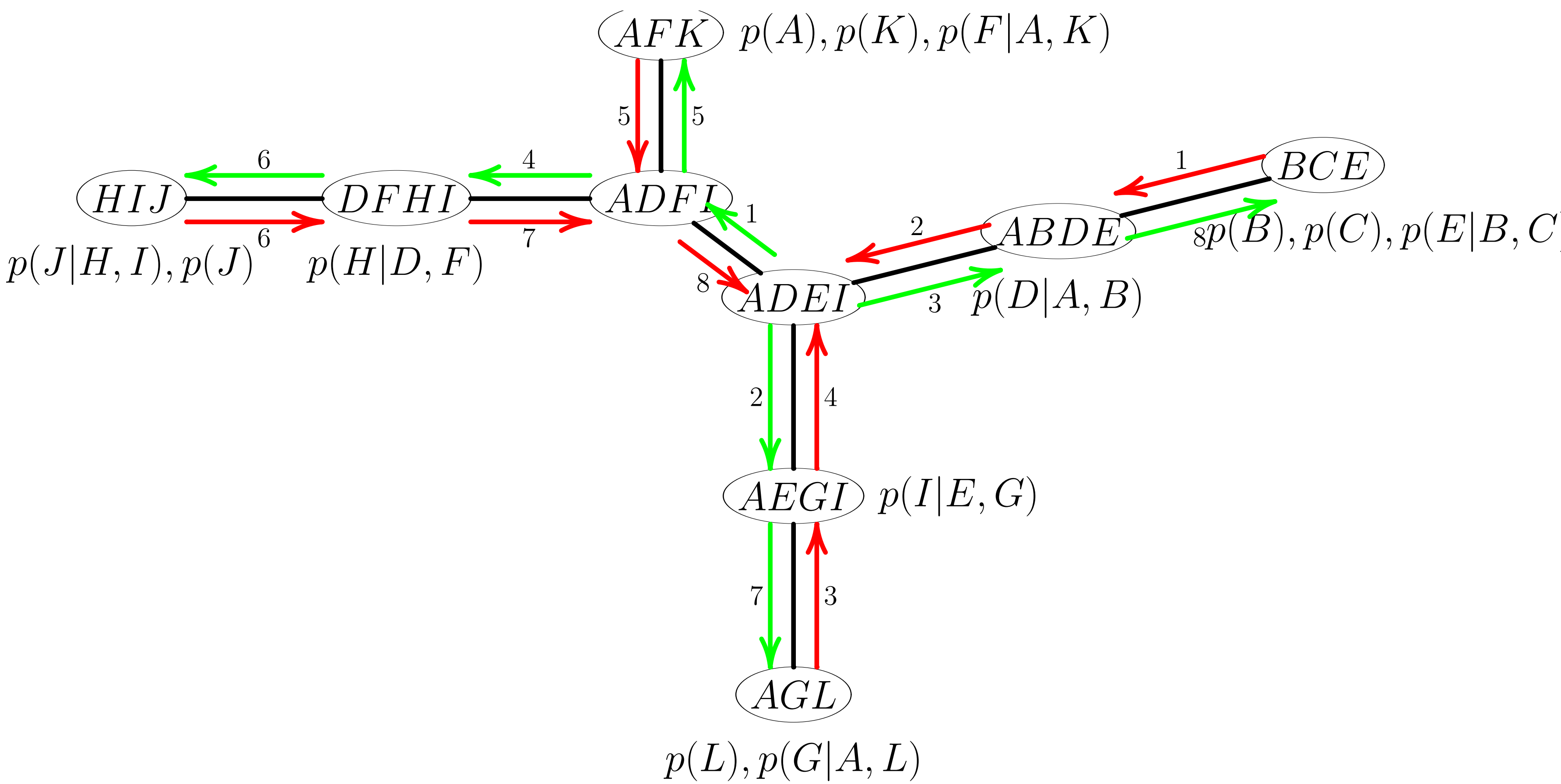
G	pure	carrier
$A =$ pure	0.0002	0.0007
carrier	0	0.0005



Studfarm / distribute evidence (8/8)

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

E	pure	carrier
$B =$ pure	0.0003	0.0009
carrier	0.0005	0.0319



Studfarm / marginalize to target domains (1/5)

marginalize to target domains:

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

$p_e(A) := p_e(AFK)^{\downarrow A} =$

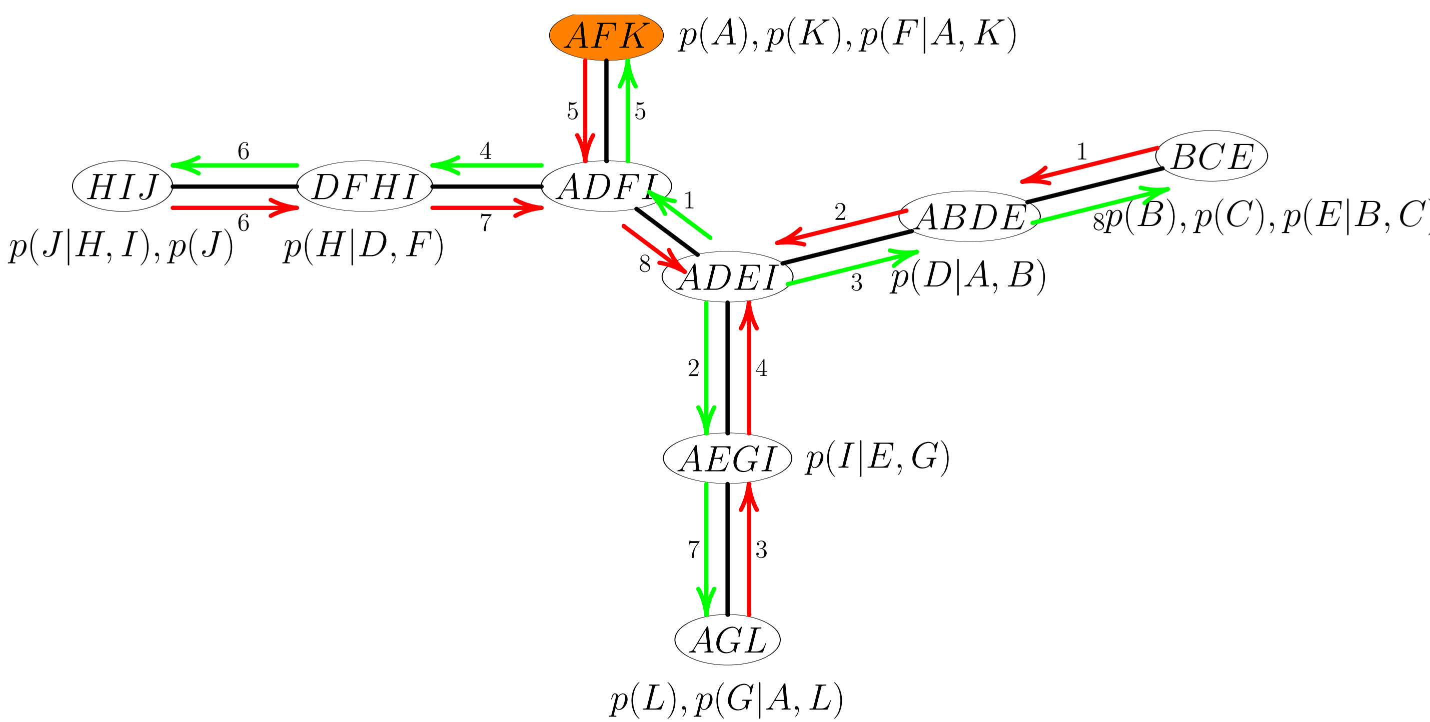
$A =$ pure	0.3764
carrier	0.6236

$p_e(F) := AFK^{\downarrow F} =$

$F =$ pure	0.5517
carrier	0.4483

$p_e(K) := AFK^{\downarrow K} =$

$K =$ pure	0.9797
carrier	0.0203



Studfarm / marginalize to target domains (2/5)

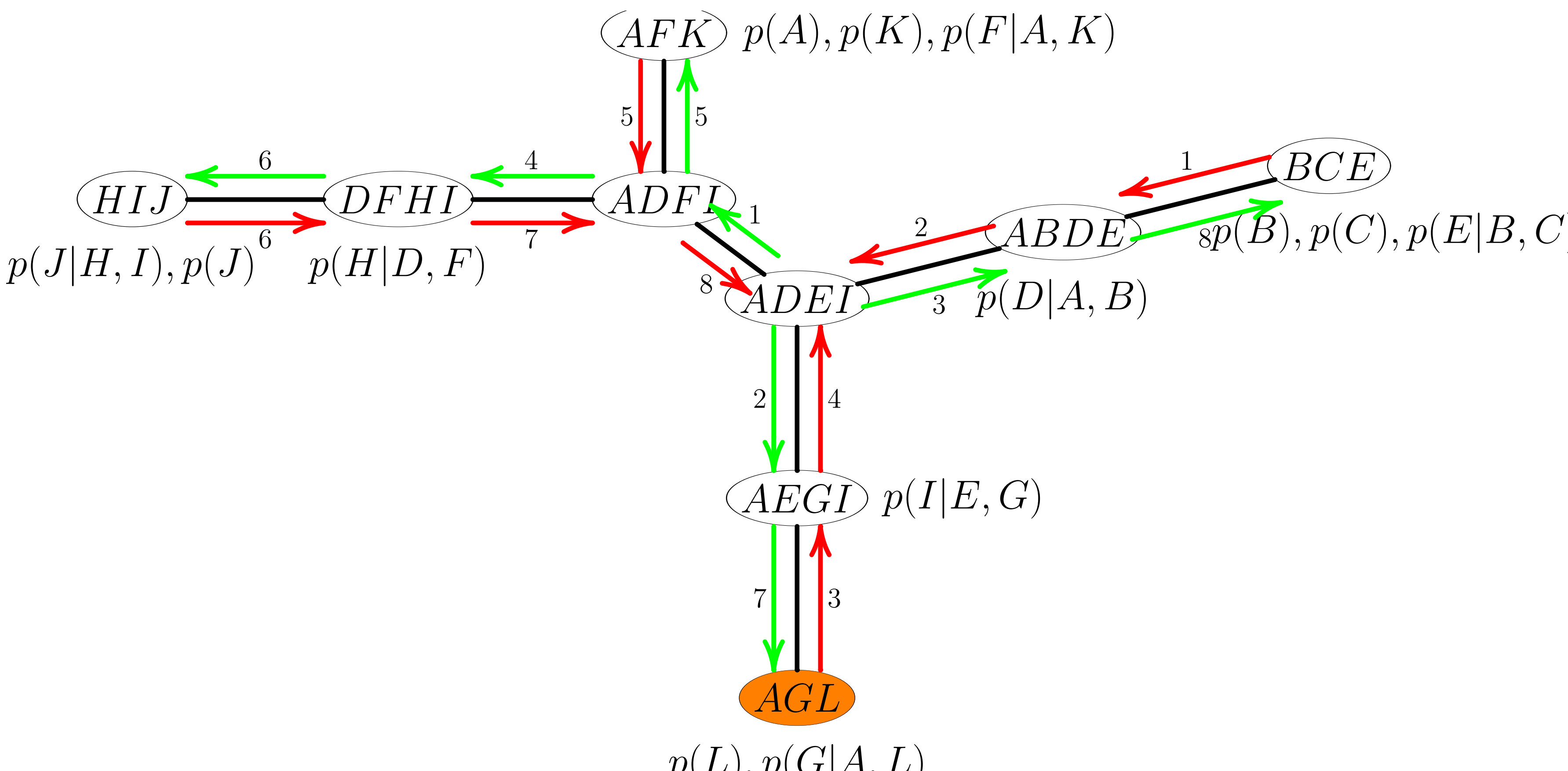
$$p_e(AGL) := q_{AEGI \to AGL} \cdot p(L) \cdot p(G|A,L)$$

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

$G =$	
pure	0.375
carrier	0.625

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

$L =$	
pure	0.982
carrier	0.018



Studfarm / marginalize to target domains (3/5)

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

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

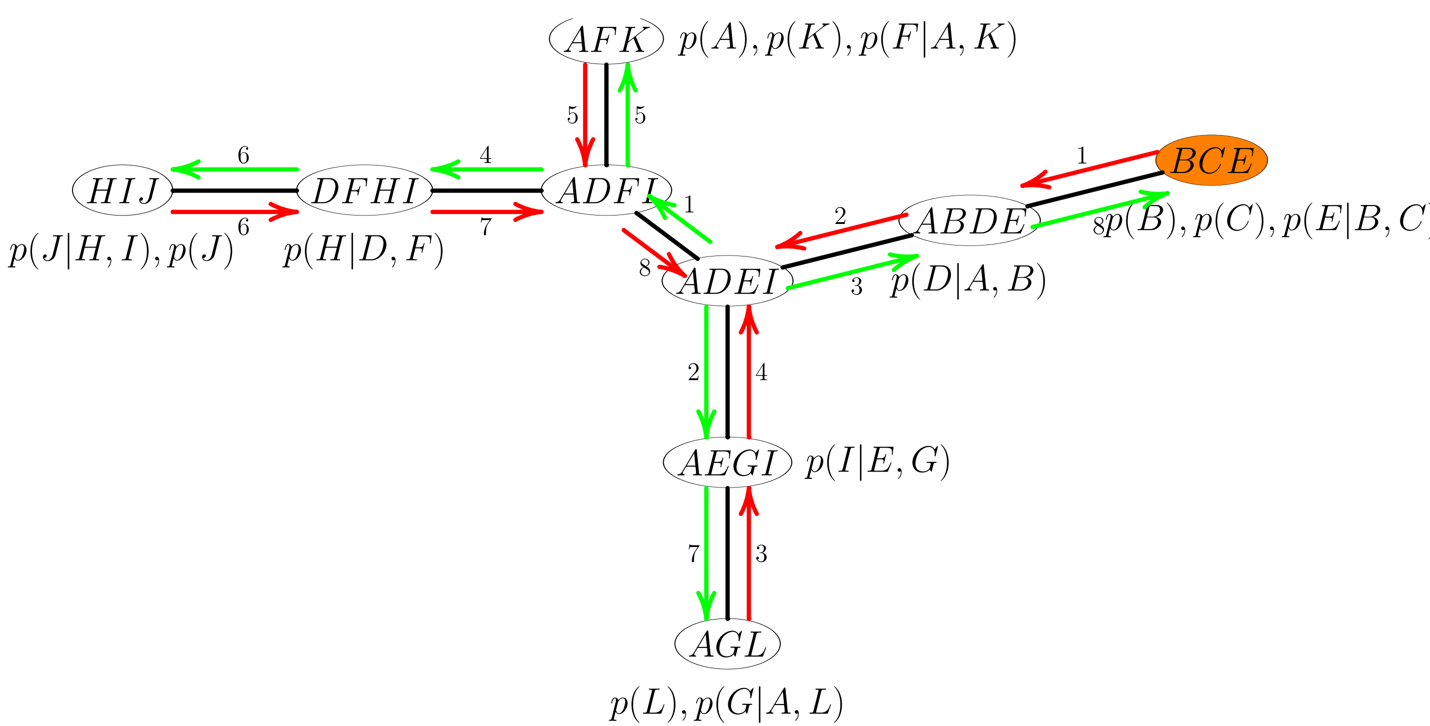
$B =$ pure	0.6192
carrier	0.3808

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

$C =$ pure	0.9812
carrier	0.0188

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

$E =$ pure	0.6138
carrier	0.3862



Studfarm / marginalize to target domains (4/5)

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

$$p_e(H) := p_e(HIJ)^{\downarrow H} =$$

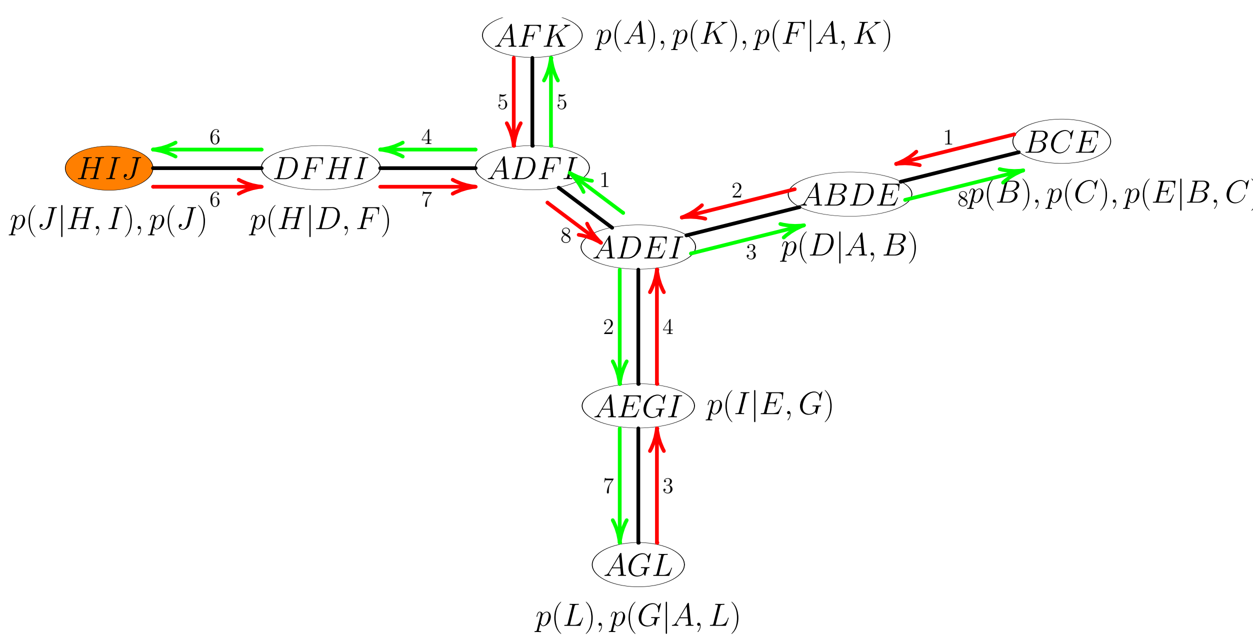
$H =$	
pure	0
carrier	1

$$p_e(I) := p_e(HIJ)^{\downarrow I} =$$

$I =$	
pure	0
carrier	1

$$p_e(J) := p_e(HIJ)^{\downarrow J} =$$

$J =$	
pure	0
carrier	0
sick	1

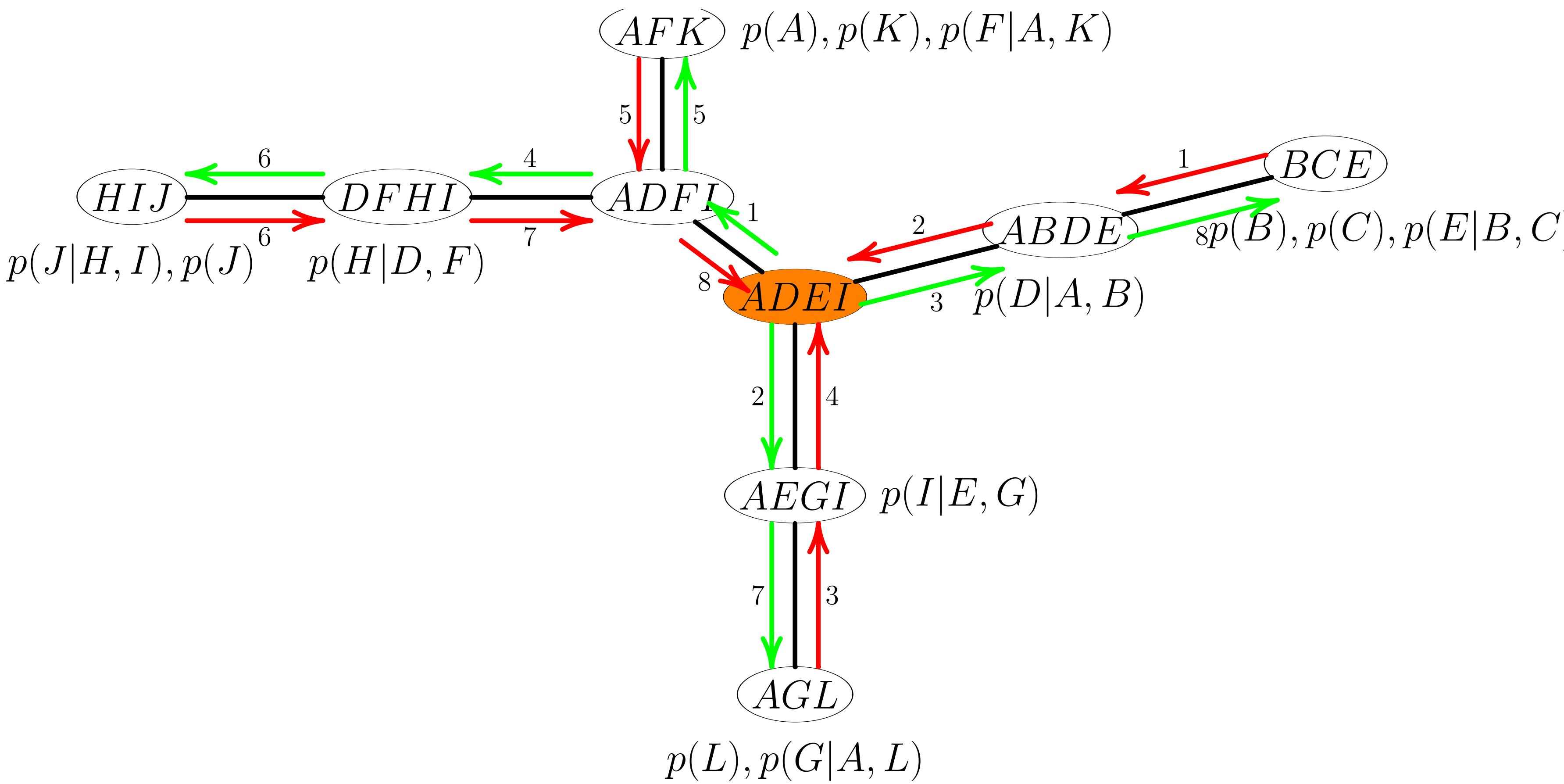


Studfarm / marginalize to target domains (5/5)

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

$$p_e(D) := p_e(ADEI)^{\downarrow D} = $$

$D =$	
pure	0.195
carrier	0.805



Studfarm / all single-variable marginals

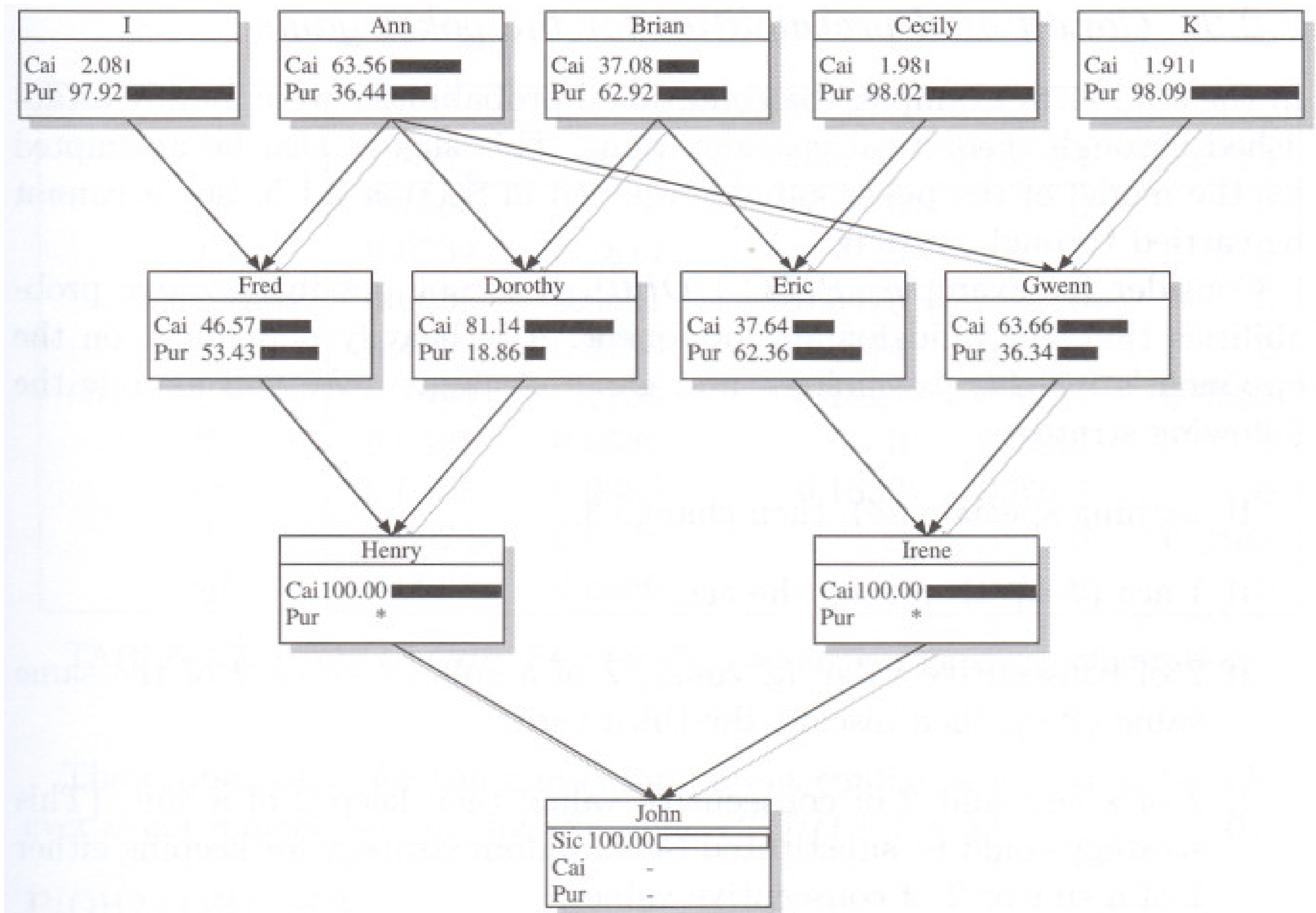


Figure 9: Probabilities given evidence that John is sick (AA). [Jen01, p. 49]

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.

We proceed as follows:

(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 $\{A, B\}$.

$p_e(A, B) =$

B	pure	carrier
$A =$ pure	0.00718	0.36919
carrier	0.61205	0.01159

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.,

$p_e(A) \cdot p_e(B) =$

$A =$ pure	0.37636
carrier	0.62364

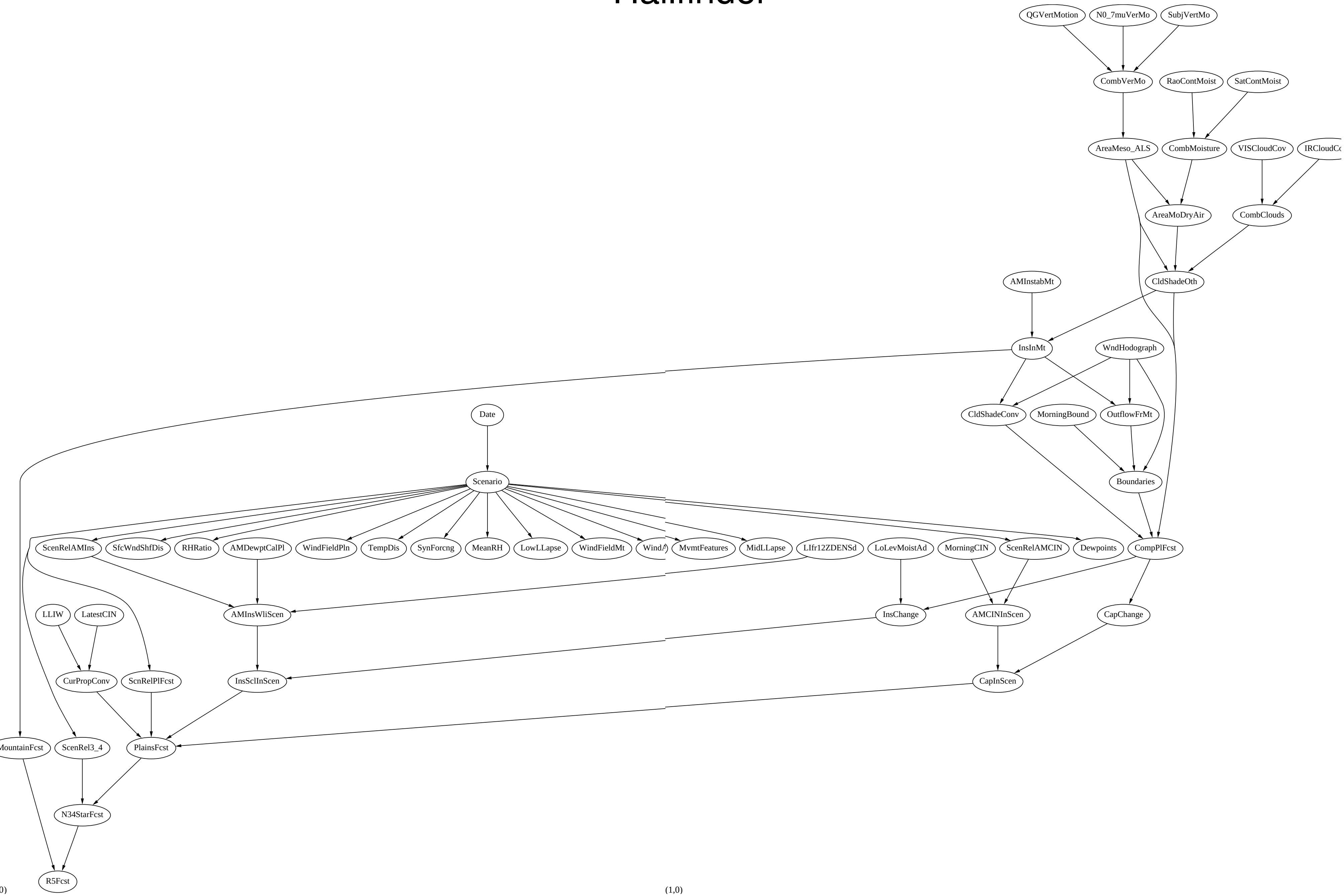
$\cdot$

$B =$ pure	0.61923
carrier	0.38077

$=$

B	pure	carrier
$A =$ pure	0.23305	0.14331
carrier	0.38617	0.23746

Hailfinder


QGVertMotion
N0_7muVerMo
SubjVertMo
CombVerMo
RaoContMoist
SatContMoist
AreaMeso_ALS
CombMoisture
VISCloudCov
IRCloudC
AreaMoDryAir
CombClouds
AMInstabMt
CldShadeOth
InsInMt
WndHodograph
Date
CldShadeConv
MorningBound
OutflowFrMt
Scenario
Boundaries
ScenRelAMIns
SfcWndShfDis
RHRatio
AMDewptCalPl
WindFieldPln
TempDis
SynForcng
MeanRH
LowLLapse
WindFieldMt
WindA
MvmtFeatures
MidLLapse
LIfr12ZDENSd
LoLevMoistAd
MorningCIN
ScenRelAMCIN
Dewpoints
CompPlFcst
LLIW
LatestCIN
AMInsWliScen
InsChange
AMCINInScen
CapChange
CurPropConv
ScnRelPlFcst
InsSclInScen
CapInScen
ountainFcst
ScenRel3_4
PlainsFcst
N34StarFcst
R5Fcst
(1,0)

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.

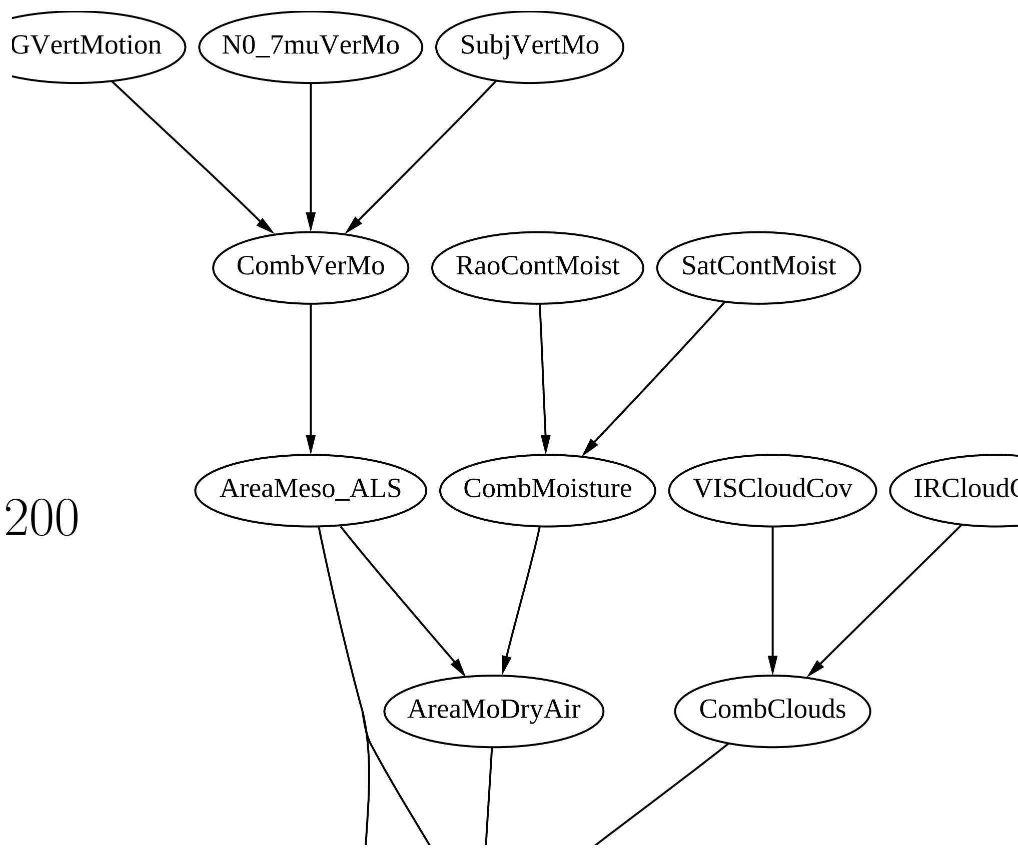


Figure 10: Hailfinder bayesian network (detail view).

Hailfinder / clique tree (minimal degree)

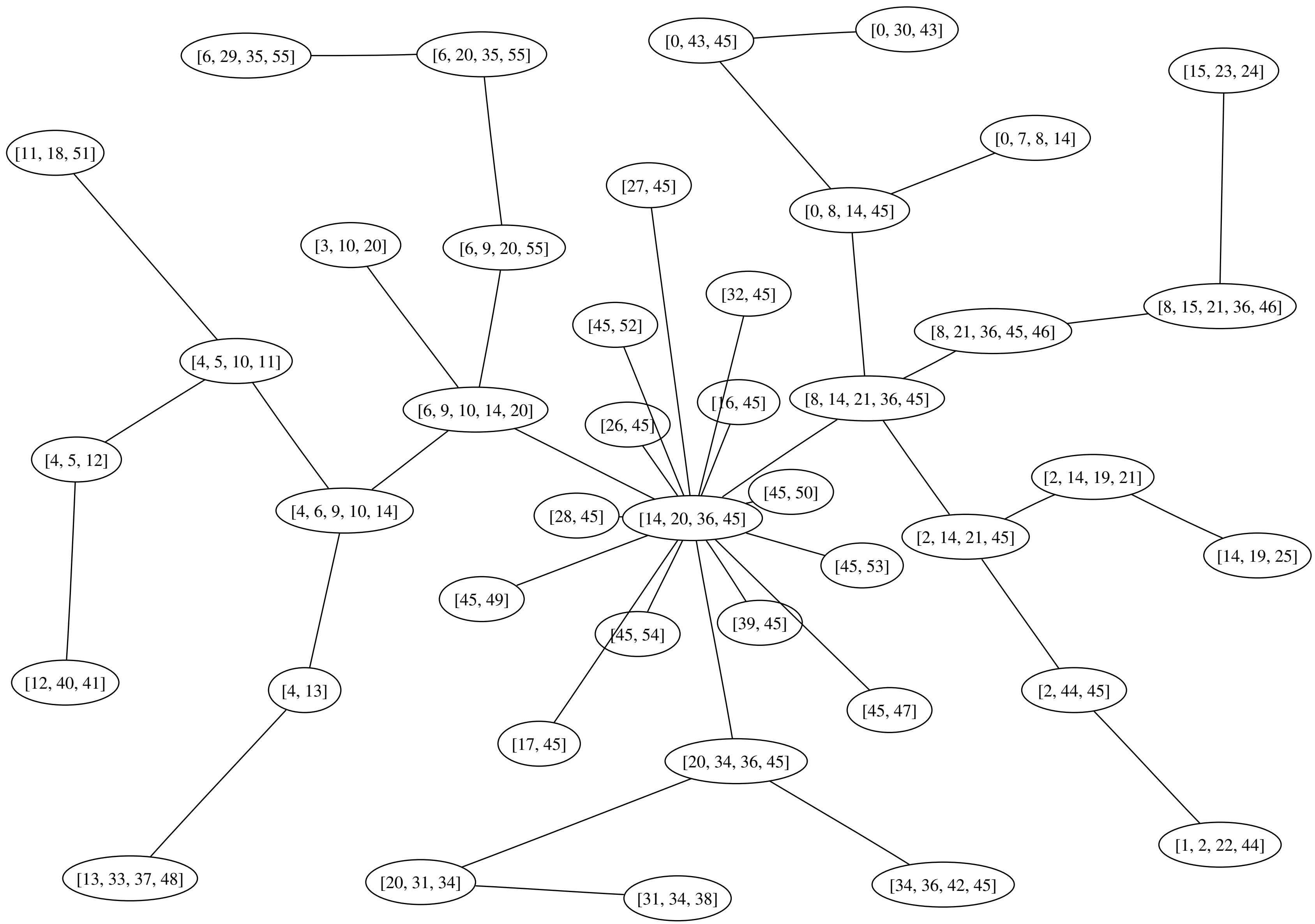


Figure 11: Clique tree of Hailfinder bayesian network.

Hailfinder / runtime

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

(i) 4s for BNJ (Lauritzen-Spiegelhalter),

(ii) 1.2s for cgnm-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.

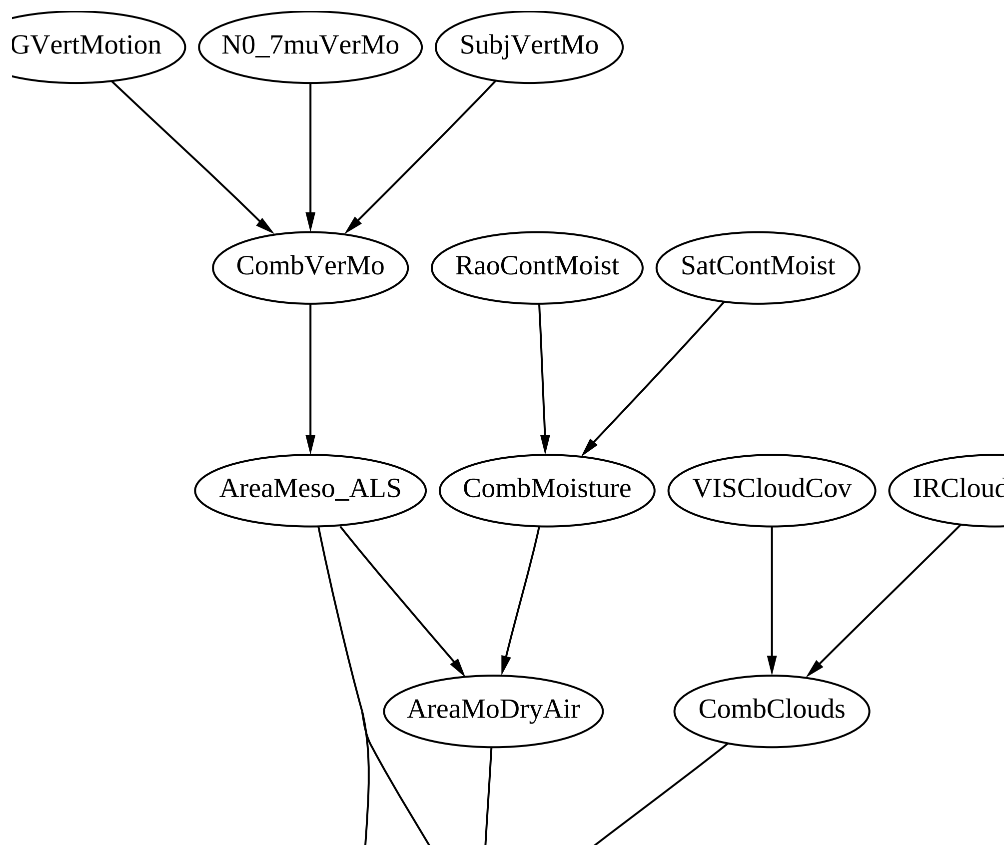


Figure 12: Hailfinder bayesian network (detail view).

References

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