NAME: Kalpana Pratapaneni (A20448916)

COLLABORATOR(S): Jasmeet(A20438656)

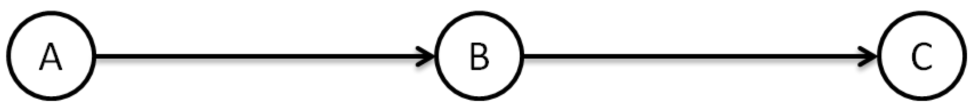
CS 583 – HOMEWORK 4

1. For the following linear chain, please calculate the requested probabilities using variable elimination. You can use any order you like. Show your work.

A	P(A)
T	0.4
F	0.6

A	B	P(B|A)
T	T	0.3
T	F	0.7
F	T	0.8
F	F	0.2

B	C	P(C|B)
T	T	0.9
T	F	0.1
F	T	0.4
F	F	0.6



a. P(C)

<u>**ANS:**</u>

Joint distribution for the above given network is = P(A)*P(B|A)*P(C|B)

Query = C, Evidence = Ø

Elimination order A, B.

So the distribution becomes $\boldsymbol{\Sigma_B P(C|B) \Sigma_A P(A)P(B|A)}$

A	B	φ(A, B) = P(A)P(B|A)
T	T	0.4 * 0.3 = 0.12
T	F	0.4 * 0.7 = 0.28
F	T	0.6 * 0.8 = 0.48
F	F	0.6 * 0.2 = 0.12

B	τ(B) = Σ_AP(A)P(B|A)
T	0.12 + 0.48 = 0.60
F	0.12 + 0.28 = 0.40

B	C	**φ(B, C) = P(C|B)τ(B)**
T	T	0.9 * 0.6 = 0.54
T	F	0.1 * 0.6 = 0.06
F	T	0.4 * 0.4 = 0.16
F	F	0.6 * 0.4 = 0.24

C	**P(C) = τ(C) = $\sum_B$φ(B, C)**
T	0.54 + 0.16 = **0.70**
F	0.06 + 0.24 = **0.30**

b. P(C|A=t)

Joint distribution for the above given network is = P(A=t)*P(B|A=t)*P(C|B)

Query = C, Evidence = A

Elimination order B.

So the distribution becomes **P(A=t)$\sum_B$P(C|B)P(B|A=t)**

B	C	**φ(B, C) = P(C|B)P(B|A=t)**
T	T	0.9 * 0.3 = 0.27
T	F	0.1 * 0.3 = 0.03
F	T	0.4 * 0.7 = 0.28
F	F	0.6 * 0.7 = 0.42

C	$\tau(C) = \Sigma_B \phi(B, C)$
T	0.27 + 0.28 = 0.55
F	0.03 + 0.42 = 0.45

C	P(C|A=t) = P(A=t)τ(C)	Normalize
T	0.4 * 0.55 = 0.22	0.22/ (0.22 + 0.18) = **0.55**
F	0.4 * 0.45 = 0.18	0.18/ (0.22 + 0.18) = **0.45**

c. P(C|A=t,B=t)

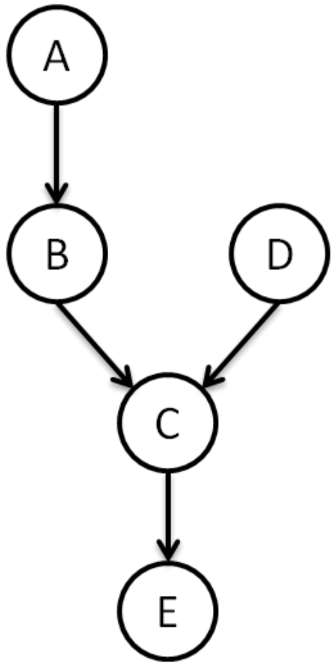
<u>Ans:</u>

Query = C, Evidence = A, B

So the distribution becomes **P(A=t)P(C|B=t)P(B=t|A=t)**

C	P(C|A=t,B=t)	Normalize
T	0.4 * 0.3 * 0.9 = 0.108	0.108 / (0.108 + 0.012) = **0.9**
F	0.4 * 0.3 * 0.1 = 0.012	0.012 / (0.108 + 0.012) = **0.1**

2. For the following Bayesian network, perform variable elimination to compute P(E). Fill in the tables. Assume the variables are binary. In the first part, use the given order. In the second part, choose an order that requires fewer operations.



a. Fill the following table for computing P(E) using the variable elimination order of B, C, A, D.
Ans:

Joint Distribution is P(A, B, C, D, E) = P(A)P(B|A)P(D)P(C|B,D)P(E|C)

Variable	All Factors	Participates	New Factor After *	# *s	New Factor After +	# +s	# Ops
B	**P(A) P(B|A) P(D) P(C|B,D) P(E|C)**	**P(B|A) P(C|B, D)**	**ϕ_1(A, B, C, D)**	**1*2*2*2 *2 = 16**	**τ_1(A, C, D)**	**1*2*2 *2 = 8**	**24**
C	**P(A) P(D) P(E|C) τ_1(A, C, D)**	**P(E|C) τ_1(A, C, D)**	**ϕ_2(A, C, D, E)**	**1*2*2*2 *2 = 16**	**τ_2(A, D, E)**	**1*2*2 *2 = 8**	**24**
A	**P(A)P(D)τ_2(A, D, E)**	**P(A)τ_2(A, D, E)**	**ϕ_3(A, D, E)**	**1*2*2*2 = 8**	**τ_3(D, E)**	**1*2*2 = 4**	**12**
D	**P(D)τ_3(D, E)**	**P(D)τ_3(D, E)**	**ϕ_4(D, E)**	**1*2*2 = 4**	**τ_4(E)**	**1*2 = 2**	**6**
Normalize	**τ_4(E)**					**1**	**3(2 divs)**
Total							**69**

b. Find a variable elimination order for computing P(E) that requires fewer computations than the part "a" above. Fill the following table for computing P(E) using that variable elimination order.

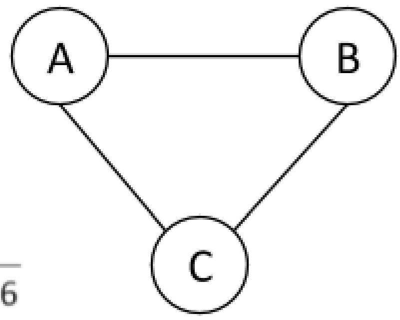
Joint Distribution is P(A, B, C, D, E) = P(A)P(B|A)P(D)P(C|B,D)P(E|C)

Variable	All Factors	Participates	New Factor After *	# *s	New Factor After +	# +s	# Ops
A	**P(A) P(B|A) P(D) P(C|B,D) P(E|C)**	**P(A) P(B|A)**	**ϕ_1(A, B)**	**1*2*2 = 4**	**τ_1(B)**	**1*2 = 2**	**6**
B	**P(D) P(C|B,D) P(E|C) τ_1(B)**	**P(C|B,D) τ_1(B)**	**ϕ_2(B, C, D)**	**1*2*2*2 = 8**	**τ_2(C, D)**	**1*2*2 = 4**	**12**
D	**P(D) P(E|C) τ_2(C, D)**	**P(D) τ_2(C, D)**	**ϕ_3(C, D)**	**1*2*2 = 4**	**τ_3(C)**	**1*2 = 2**	**6**
C	**P(E|C) τ_3(C)**	**P(E|C) τ_3(C)**	**ϕ_4(C, E)**	**1*2*2 = 4**	**τ_4(E)**	**1*2 = 2**	**6**
Normalize	**τ_4(E)**					**1**	**3(2 divs)**
Total							**33**

3. For the following Markov network, calculate the requested probabilities using variable elimination. You can use any order you like.

A	B	ϕ(A,B)
T	T	5
T	F	1
F	T	1
F	F	5

A	ϕ(A)
T	2
F	1



B	ϕ(B)
T	1
F	4

A	C	ϕ(A,C)
T	T	6
T	F	1
F	T	1
F	F	6

C	ϕ(C)
T	1
F	8

B	C	ϕ(B,C)
T	T	1
T	F	10
F	T	10
F	F	1

a. P(A)

Markov distribution for the given network is,

p(A, B, C) ∝ ϕ(A)ϕ(B)ϕ(C)ϕ(A, B)ϕ(B, C)ϕ(C, A)

Query Y= A, Evidence is empty. Let's consider the elimination order as B, C.

Eliminate B, then **φ**(A, B, C) = ϕ(B) * ϕ(A, B) * ϕ(B, C)

A	B	C	**φ(A, B, C)**
T	**T**	**T**	**1 * 5 * 1 = 5**
T	**T**	**F**	**1 * 5 * 10 = 50**
T	**F**	**T**	**4 * 1 * 10 = 40**
T	**F**	**F**	**4 * 1 * 1 = 4**
F	**T**	**T**	**1 * 1 * 1 = 1**
F	**T**	**F**	**1 * 1 * 10 = 10**

F	F	T	4 * 5 * 10 = 200
F	F	F	4 * 5 * 1 = 20

Eliminate B, then ⊤(A, B, C) = $\sum_B$(A, C)

A	C	⊤(A, C)
T	T	5 + 40 = 45
T	F	50 + 4 = 54
F	T	1 + 200 = 201
F	F	10 + 20 = 30

And the distribution becomes

ϕ(A)ϕ(C)ϕ(C, A)⊤(A, C)

Eliminating C, hen **φ**(A, C) = ϕ(C) * ϕ(C, A) * ⊤(A, C)

A	C	φ(A, C)
T	T	1 * 6 * 45 = 270
T	F	8 * 1 * 54 = 432
F	T	1 * 1 * 201 = 201
F	F	8 * 6 * 30 = 1440

A	⊤(A)
T	270 + 432 = 702
F	201 + 1440 = 1641

And the distribution becomes ϕ(A)⊤(A)

A	$⊤_2$(A) = ϕ(A)⊤(A)	P(A)
T	2 * 702 = 1404	1404 / (1404 + 1641) = 0.46
F	1 * 1641 = 1641	1641 / (1404 + 1641) = 0.54

b. P(A|B=t)

Markov distribution for the given network is,

p(A, B, C) $\propto$ ϕ(A)ϕ(B)ϕ(C)ϕ(A, B)ϕ(B, C)ϕ(C, A)

Given Query Y= A, Evidence isB=t, elimination order is C.

ϕ(A)ϕ(B=t)ϕ(C)ϕ(A, B=t)ϕ(B=t, C)ϕ(C, A)

φ(A, B, C) = ϕ(C)ϕ(B=t, C)ϕ(C, A)

A	C	φ(A, B=t, C)
T	T	1 * 1 * 6 = 6
T	F	8 * 10 * 1 = 80
F	T	1 * 1 * 1 = 1
F	F	8 * 10 * 6 = 480

A	⊤(A, B=t)
T	6 + 80 = 86
F	1 + 480 = 481

Finally $\top_2$(A, B=t) = ϕ(A)ϕ(B=t)ϕ(A, B=t)⊤(A, B=t)

A	⊤(A, B=t)	P(A|B=t)
T	2 * 1 * 5 * 86 = 860	860/(860+481) = 0.64
F	1 * 1 *1 * 481 = 481	481/(860+481) = 0.36

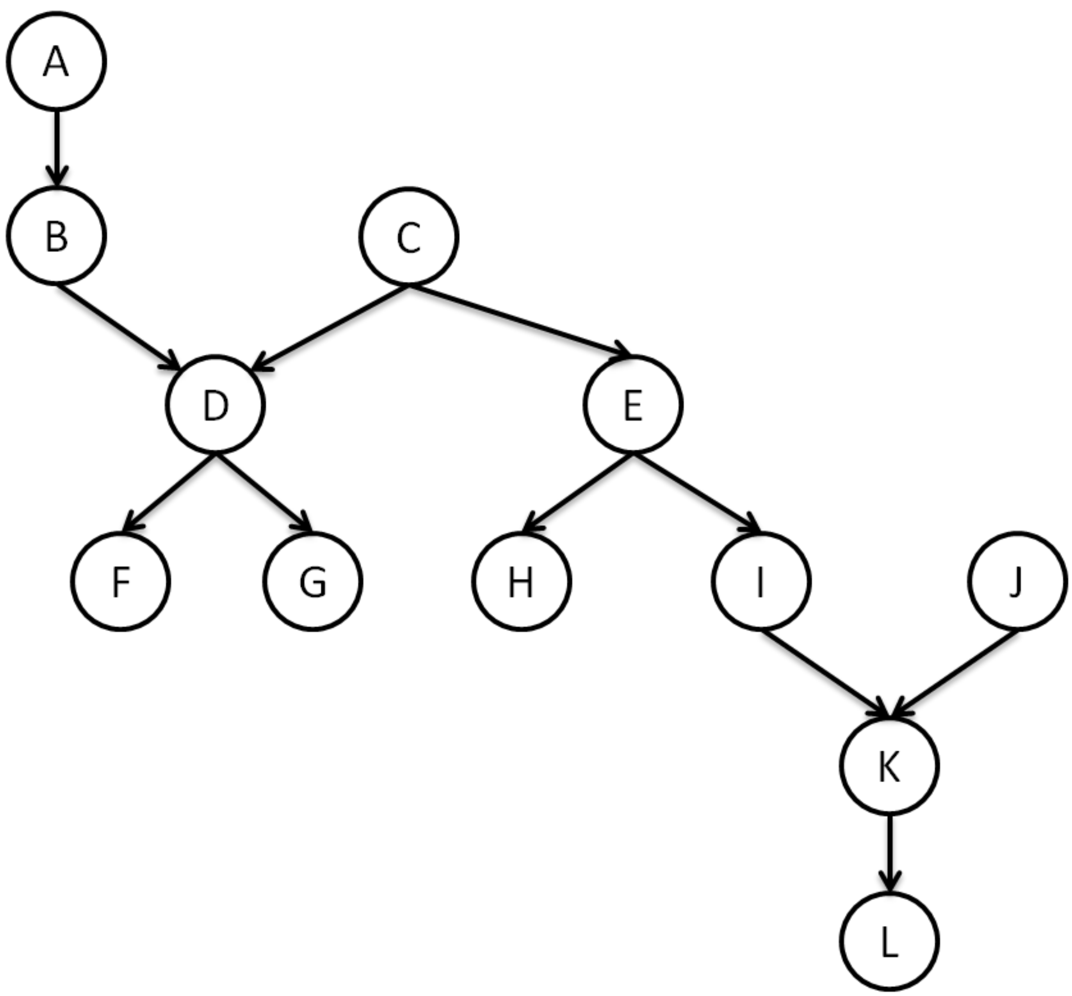
c. P(A|B=t, C=f)

Given Query Y= A, Evidence isB=t, C=f.

P(A|B=t, C=f) $\propto$ ϕ(A)ϕ(B=t)ϕ(C=f)ϕ(A, B=t)ϕ(B=t, C=f)ϕ(C=f, A)

A	τ(A, B=t, C=f)	P(A|B=t, C=f)
T	2 * 1 * 8 * 5 * 10 * 1 = 800	800 / (800 + 480) = 0.625
F	1 * 1 * 8 * 1 * 10 * 6 = 480	480 / (800 + 480) = 0.375

4. We are given the following Bayesian network. For each of the following queries, indicate which variables are irrelevant.



Joint Distribution for the above network is,

P(A, B, C, D, E, F, G, H, I, J, K, L) = P(A) P(B|A) P(C) P(D|B, C) P(F|D) P(G|D) P(E|C) P(H|E) P(I|E) P(J) P(K|I, J) P(L|K)

a. P(E)

A, B, D, F, G, H, I, J, K, L

b. P(E|K)

A, B, D, F, G, H, L

c. P(E|L)

A, B, D, F, G, H

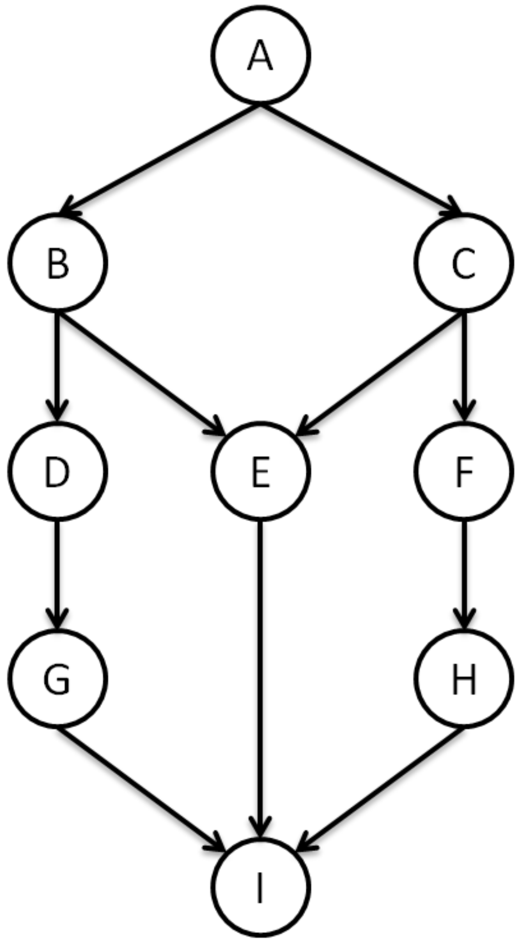
d. P(E|G)

F, H, I, J, K, L

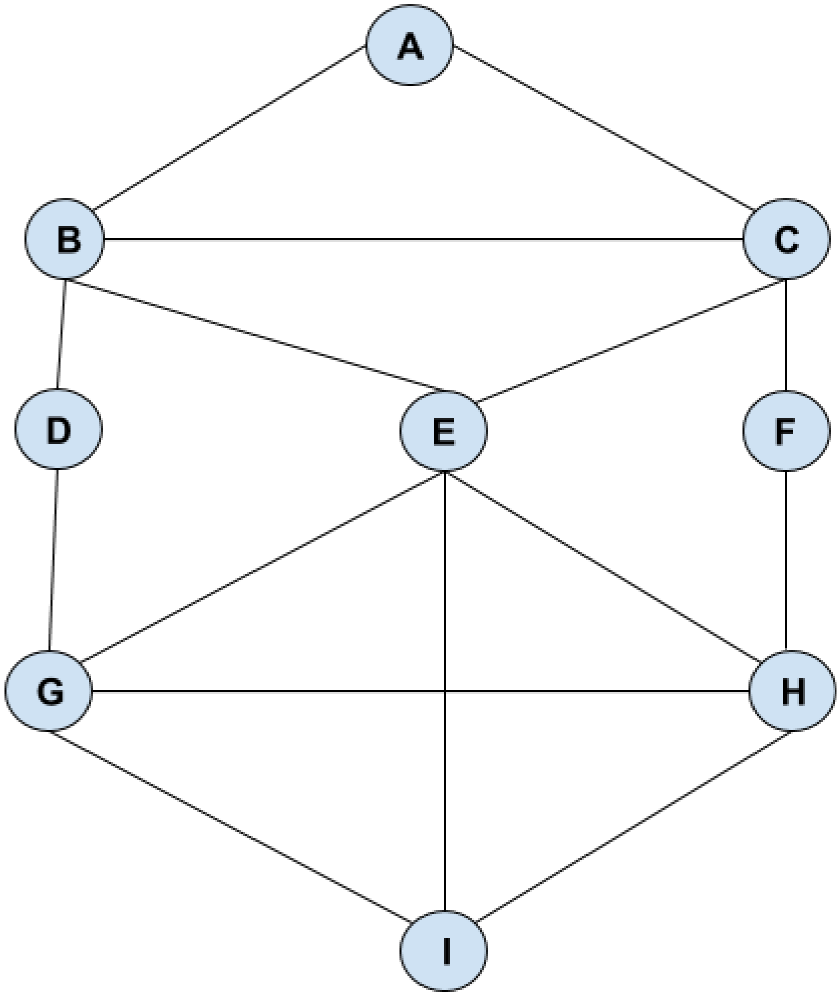
e. P(E|G, J)

F, H, I, K, L

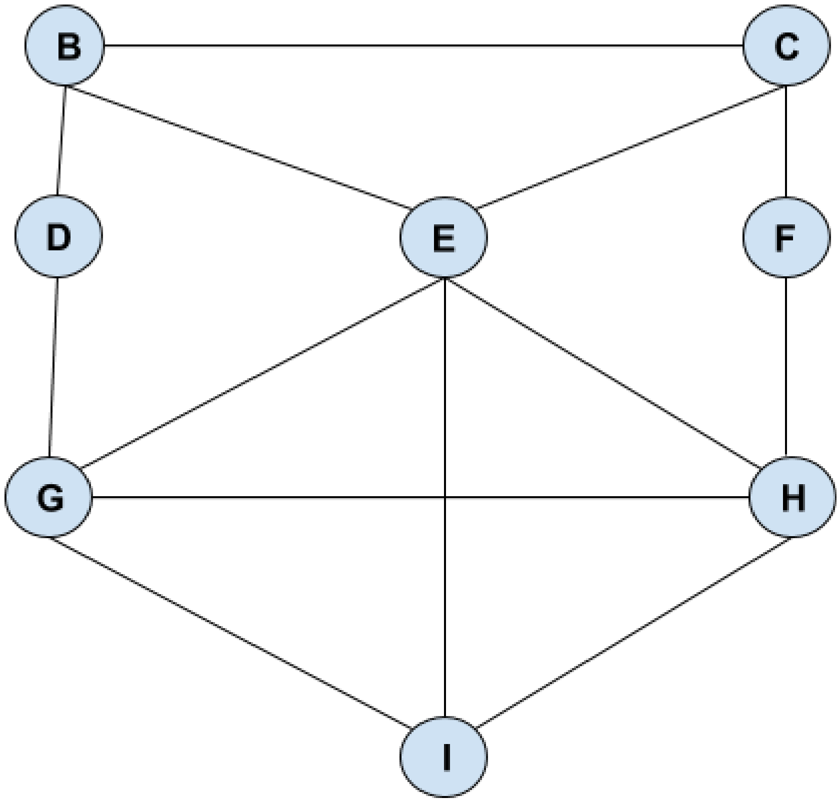
5. For the following Bayesian network, create a join tree. First, moralize the graph. Then, pick an order using the min-fill heuristic. After finding the maximal cliques, connect those maximal cliques using maximal sepsets. Make sure that the final tree is a clique tree; that is, it is family preserving and it has the running intersection property.



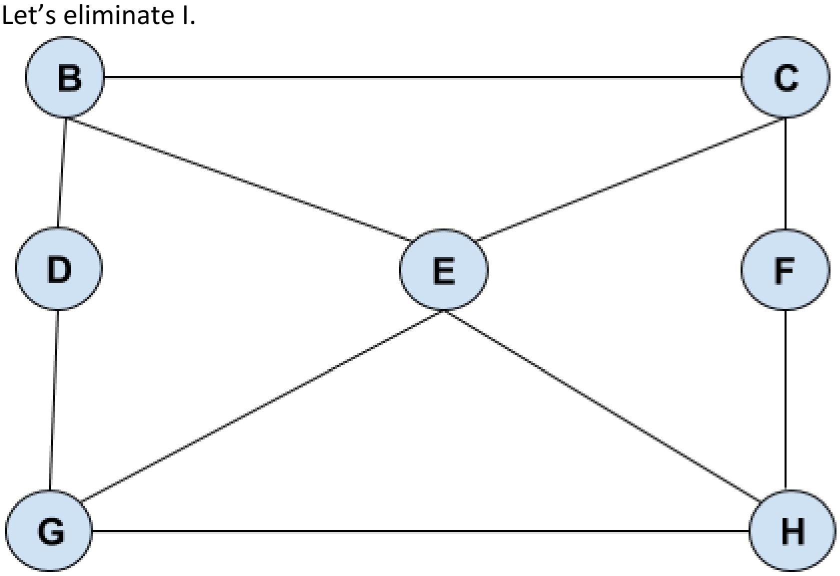
Moralizing the graph:



Based on min-fill heuristic, let's eliminate Node A.

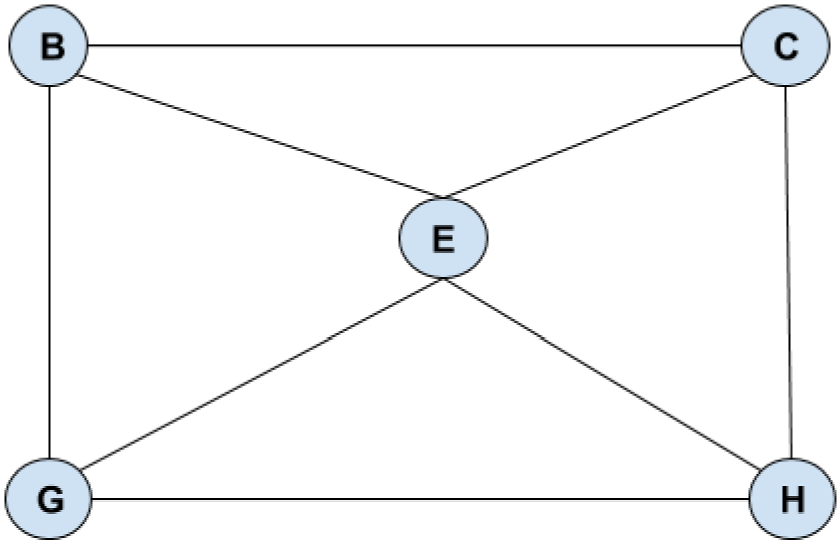


A’s edges B and C are already connected, so we are not introducing a new edge.
Maximal Cliques: ABC.
Let’s eliminate I.

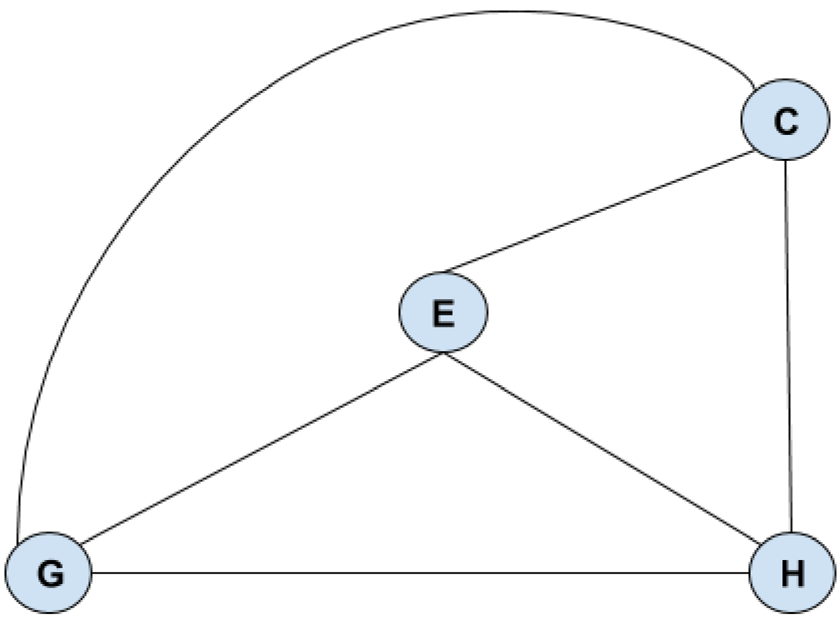


Maximal Cliques: ABC, EGHI.

Let’s eliminate D and F.

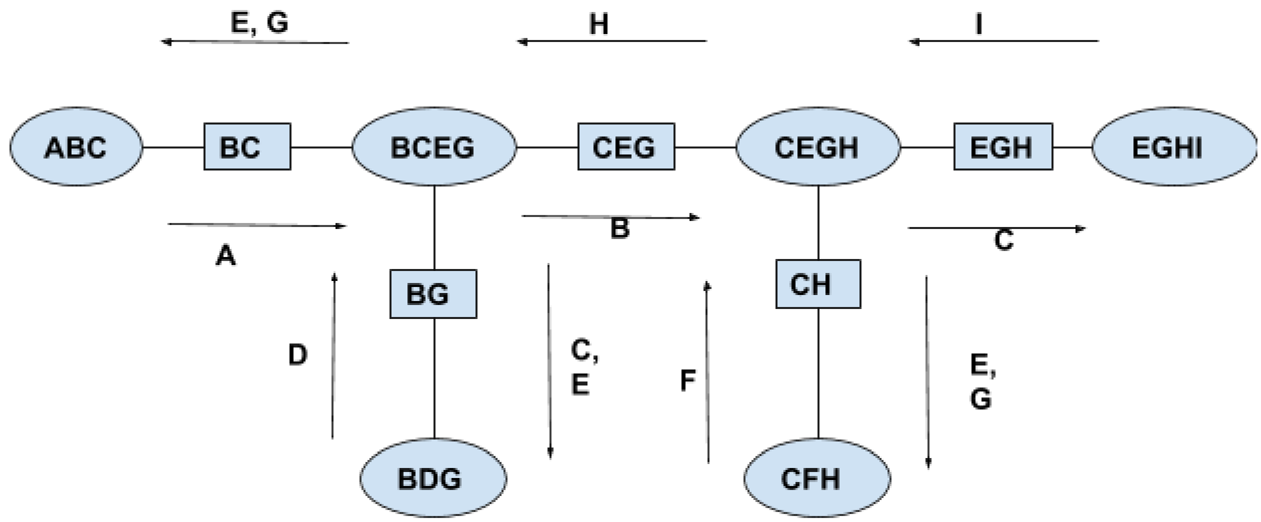


Maximal Cliques: ABC, EGHI, BDG, CFH.
Let’s eliminate G.

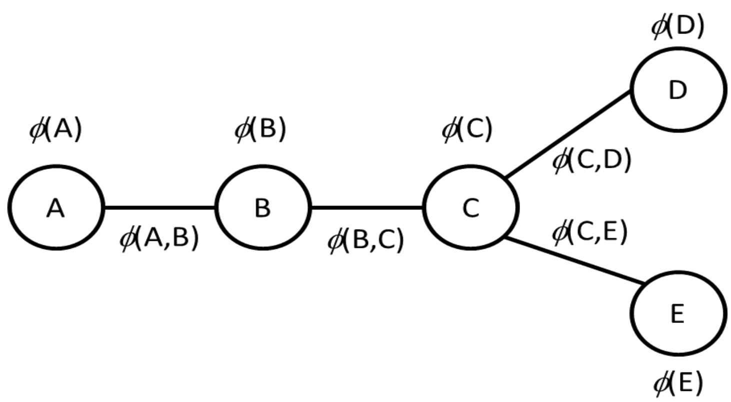


Maximal Cliques: ABC, EGHI, BDG, CFH, BCEG.
Let’s eliminate C, E, G and H.

Finals maximal cliques are: **ABC, EGHI, BDG, CFH, BCEG, CEGH**



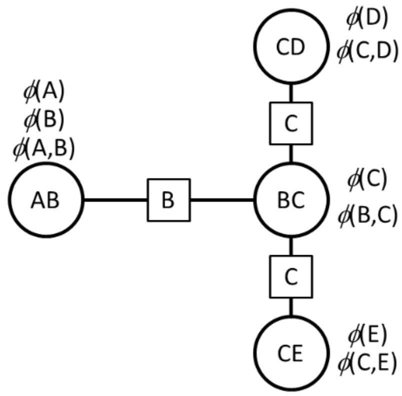
6. We are given the following pair-wise Markov random field.



All variables are binary. The potentials are defined as follows.

A prefers to be True: $\phi(A) = \langle 2, 1\rangle$. B and C do not have any preferences: $\phi(B) = \phi(C) = \langle 1, 1\rangle$. D and prefer to be False: $\phi(D) = \phi(E) = \langle 1, 2\rangle$. The pair-wise potentials prefer positive correlations. $\phi(X_i, X_j) = 4$ if $X_i = X_j$, 1 otherwise, for $X_i, X_j \in$ {A, B, C, D}. For example, $\phi(A{=}T, B{=}T) = \phi(A{=}F, B{=}F) = 4$ and $\phi(A{=}T, B{=}F) = \phi(A{=}F, B{=}T) = 1$.

Here is a junction tree for this network where potentials are assigned to their cluster nodes.



a. Perform message-passing on this junction tree. Compute all the messages. That is, compute

i. $\delta_{AB \to BC}$

$\delta_{AB \to BC} = \Sigma_A \phi(A)\, \phi(B)\, \phi(A, B)$

A	B	φ(A, B) = ϕ(A) ϕ(B) ϕ(A, B)
T	**T**	**2 * 1 * 4 = 8**
T	**F**	**2 * 1 * 1 = 2**
F	**T**	**1 * 1 * 1 = 1**
F	**F**	**1 * 1 * 4 = 4**

B	τ_1(B)
T	**8 + 1 = 9**
F	**2 + 4 = 6**

$\delta_{AB \to BC =}\ \tau_1$(B) = <9, 6>

ii. $\delta_{BC \to CD}$

$\delta_{BC \to CD} = \Sigma_B \phi(C)\, \phi(B, C) * \delta_{AB \to BC} * \delta_{CE \to BC}$

$\delta_{CE \to BC}$ = get it from 6th - v answer.

B	C	φ(B, C) = ϕ(C) ϕ(B, C) * τ_1(B) * τ_5(C)
T	**T**	**1 * 4 * 9 * 6 = 216**
T	**F**	**1 * 1 * 9 * 9 = 81**

F	T	1 * 1 * 6 * 6 = 36
F	F	1 * 4 * 6 * 9 = 216

C	τ_2(C)
T	216 + 36 = 252
F	81 + 216 = 297

iii. $\delta_{BC \to CE}$

$\delta_{BC\to CE} = \Sigma_B\, \phi(C) * \phi(B, C) * \delta_{AB\to BC\,*}\, \delta_{CD\to BC}$

B	C	$\varphi(C, D) = \phi(C)\, \phi(B, C) * \tau_1(B) * \tau_4(C)$
T	T	1 * 4 * 9 * 6 = 216
T	F	1 * 1 * 9 * 9 = 81
F	T	1 * 1 * 6 * 6 = 36
T	F	1 * 4 * 6 * 9 = 216

C	τ_3(C)
T	216 + 36 = 252
F	81 + 216 = 297

iv. $\delta_{CD \rightarrow BC}$

$\delta_{CD \rightarrow BC} = \Sigma_D \phi(D) \phi(C, D)$

C	D	$\varphi(C, D) = \phi(D) \phi(C, D)$
T	T	1 * 4 = 4
T	F	2 * 1 = 2
F	T	1 * 1 = 1
T	F	2 * 4 = 8

C	$\tau_4(C)$
T	4 + 2 = 6
F	1 + 8 = 9

v. $\delta_{CE \rightarrow BC}$

$\delta_{CE \rightarrow BC} = \Sigma_E \phi(E) * \phi(C, E)$

C	E	$\varphi(C, E) = \phi(E) \phi(C, E)$
T	T	1 * 4 = 4
T	F	2 * 1 = 2
F	T	1 * 1 = 1
F	F	2 * 4 = 8

C	$\tau_5(C)$
T	6
F	9

vi. $\delta_{BC \to AB}$

$\delta_{BC \to AB} = \Sigma_C\, \phi(C)\, \phi(B, C) * \delta_{CD \to BC} * \delta_{CE \to BC}$

B	C	$\varphi(C, D) = \phi(C)\, \phi(B, C) * \tau_4(C) * \tau_5(C)$
T	T	1 * 4 * 6 * 6 = 144
T	F	1 * 1 * 9 * 9 = 81
F	T	1 * 1 * 6 * 6 = 36
T	F	1 * 4 * 9 * 9 = 324

B	$\tau_6(B)$
T	144 + 81 = 225
F	36 + 324 = 360

b. What is the marginal probability of each individual variable? That is, compute the following using the potentials assigned to cluster nodes and messages those cluster nodes received from their neighbors

i. P(A)

Let's pick the node AB. calculate Belief at AB i.e B_{AB}

$B_{AB} = \phi(A)\, \phi(B)\, \phi(A, B)\ \delta_{BC \to AB}$

$= \phi(A) * \phi(B) * \phi(A, B) * \tau_6(B)$

A	B	$\varphi(A, B) = \phi(A) * \phi(B) * \phi(A, B) * \tau_6(B)$
T	T	2 * 1 * 4 * 225 = 1800
T	F	2 * 1 * 1 * 360 = 720
F	T	1 * 1 * 1 * 225 = 225
F	F	1 * 1 * 4 * 360 = 1440

A	τ_7(A)	P(A)
T	2520	2520 / (2520 + 1665) = 0.6
F	1665	1665 / (2520 + 1665) = 0.4

ii. P(B)

$B_{AB} = \phi(A)\,\phi(B)\,\phi(A, B)\ \delta_{BC \to AB}$

$= \phi(A) * \phi(B) * \phi(A, B) * \tau_6(B)$

A	B	$\varphi(A, B) = \phi(A) * \phi(B) * \phi(A, B) * \tau_6(B)$
T	T	2 * 1 * 4 * 225 = 1800
T	F	2 * 1 * 1 * 360 = 720
F	T	1 * 1 * 1 * 225 = 225
F	F	1 * 1 * 4 * 360 = 1440

B	τ_8(B)	P(B)
T	1800 + 225 = 2025	2025 / (2025 + 2160) = 0.48
F	720 + 1440 = 2160	2160 / (2025 + 2160) = 0.52

iii. P(C)

$B_{CD} = \phi(D) * \phi(C, D) * \delta_{BC \to CD}$

$= \phi(D) * \phi(C, D) * \tau_2(C)$

C	D	$\varphi(C, D) = \phi(D) * \phi(C, D) * \tau_2(C)$
T	T	1 * 4* 252 = 1008
T	F	2 * 1* 252 = 504
F	T	1 * 1* 297 = 297

F	F	2 * 4* 297 = 2376

C	τ_9(C)	P(C)
T	1008 + 504 = 1512	1512 / (1512 + 2673) = 0.36
F	297 + 2376 = 2673	2673 / (1512 + 2673) = 0.64

iv. P(D)

$B_{CD} = \phi(D) * \phi(C, D) * \delta_{BC \rightarrow CD}$

$= \phi(D) * \phi(C, D) * \tau_2(C)$

C	D	$\varphi(C, D) = \phi(D) * \phi(C, D) * \tau_2(C)$
T	T	1 * 4* 252 = 1008
T	F	2 * 1* 252 = 504
F	T	1 * 1* 297 = 297
F	F	2 * 4* 297 = 2376

D	τ_9(D)	P(D)
T	1008 + 297 = 1305	1305 / (1305 + 2880) = 0.31
F	504 + 2376 = 2880	2880 / (1305 + 2880) = 0.69

v. P(E)

B_{CE} = ϕ(E) * ϕ(C, E) * $\delta_{BC \rightarrow CE}$

= ϕ(E) * ϕ(C, E) * τ_3(C)

C	E	φ(C, E) = ϕ(E) * ϕ(C, E) * τ_2(C)
T	T	1 * 4 * 252 = 1008
T	F	2 * 1 * 252 = 504
F	T	1 * 1 * 297 = 297
F	F	2 * 4 * 297 = 2376

E	τ_{10}(E)	P(E)
T	1008 + 297 = 1305	1305 / (1305 + 2880) = 0.31
F	504 + 2376 = 2880	2880 / (1305 + 2880) = 0.69