NAME: Kalpana Pratapaneni
A* : A20448916
COLLABORATOR(S): Tushar Nitave, Jasmeet

CS 583 – Assignment 5

For all problems, please use the random numbers provided in random.csv file.

1. **Forward Sampling**: For the student network provided in class, assume that the evidence set is empty. We would like to compute marginal probabilities. We can of course perform variable elimination on this network because it is small, but for the purposes of this exercise, we will perform sampling. Perform forward sampling, assuming the variable order of D, I, S, G, L. For random number generator, use the random numbers provided in the random.csv file.
   a. Complete the following table. Continue sampling until you run out of random numbers. Add rows as needed. The first row is given as an example.

D	**I**	**S**	**G**	**L**
r=0.489 P(D)=<0.6,0.4> sampled: $\mathbf{d^0}$	r=0.833 P(I)=<0.7,0.3> sampled: $\mathbf{i^1}$	r=0.869 $P(S \mid i^1)$=<0.2,0.8> sampled: $\mathbf{s^1}$	r=0.652 $P(G \mid d^0,i^1)$=<0.9,0.08,0.02> sampled: $\mathbf{g^1}$	r=0.581 $P(L \mid g^1)$=<0.1,0.9> sampled: $\mathbf{l^1}$
r = 0.042 P(D)=<0.6,0.4> sampled: $\mathbf{d^0}$	r = 0.748 P(I)=<0.7,0.3> Sampled: $\mathbf{i^1}$	r = 0.549 $P(S \mid i^1)$=<0.2,0.8> Sampled: $\mathbf{s^1}$	r = 0.49 $P(G \mid d^0,i^1)$=<0.9,0.08,0.02> Sampled: $\mathbf{g^1}$	r = 0.023 $P(L \mid g^1)$=<0.1,0.9> Sampled: $\mathbf{l^0}$
r = 0.849 P(D)=<0.6,0.4> sampled: $\mathbf{d^1}$	r = 0.009 P(I)=<0.7,0.3> Sampled: $\mathbf{i^0}$	r = 0.612 $P(S \mid i^0)$ = <0.95, 0.05> sampled: $\mathbf{s^0}$	r = 0.936 $P(G \mid d^1,i^0)$=<0.05,0.25,0.7> sampled: $\mathbf{g^3}$	r = 0.417 $P(l \mid g3)$=<0.99, 0.01> sampled: $\mathbf{l^0}$
r = 0.63 P(D)=<0.6,0.4> sampled: $\mathbf{d^1}$	r = 0 P(I)=<0.7,0.3> sampled: $\mathbf{i^0}$	r = 0.873 $P(S \mid i^0)$ = <0.95, 0.05> sampled: $\mathbf{s^0}$	r = 0.453 $P(G \mid d^1,i^0)$=<0.05,0.25,0.7> sampled: $\mathbf{g^3}$	r = 0.89 $P(l \mid g3)$=<0.99, 0.01> sampled: $\mathbf{l^0}$
r = 0.768 P(D)=<0.6,0.4> sampled: $\mathbf{d^1}$	r = 0.859 P(I)=<0.7,0.3> sampled: $\mathbf{i^1}$	r = 0.394 $P(S \mid i^1)$=<0.2,0.8> sampled: $\mathbf{s^1}$	r = 0.157 $P(G \mid d^1,i^1)$=<0.5,0.3,0.2> sampled: $\mathbf{g^1}$	r = 0.103 $P(L \mid g^1)$=<0.1,0.9> Sampled: $\mathbf{l^1}$
r = 0.81 P(D)=<0.6,0.4> sampled: $\mathbf{d^1}$	r = 0.139 P(I)=<0.7,0.3> sampled: $\mathbf{i^0}$	r = 0.037 $P(S \mid i^0)$ = <0.95, 0.05> sampled: $\mathbf{s^0}$	r = 0.12 $P(G \mid d^1,i^0)$=<0.05,0.25,0.7> sampled: $\mathbf{g^2}$	r = 0.51 $P(L \mid g^1)$=<0.1,0.9> Sampled: $\mathbf{l^1}$

b. What is P(I) with respect to your sample?

I	Count	Normalize
i^0	3	0.5
i^1	3	0.5

P(I) = <0.5, 0.5>

2. Rejection Sampling: Assume the evidence is $S=s^0$, $G=g^3$. Perform rejection sampling on the network. When an instance is guaranteed to be rejected, do not waste any more resources on that instance; move to the next instance immediately. If an instance is rejected, note it as so.

a. **Complete the following table. Continue sampling until you run out of random numbers. Add rows as needed. The first two rows are given as examples.**

D	I	$S=s^0$	$G=g^3$	L	Rej ?
r=0.489 **P(D)=<0.6,0.4>** **sampled: d^0**	**r=0.833** **P(I)=<0.7,0.3>** **sampled: i^1**	**r=0.869** **$P(S\|i^1)$=<0.2,0.8>** **sampled: s^1**	---	---	**Yes**
r=0.652 **P(D)=<0.6,0.4>** **sampled: d^1**	**r=0.581** **P(I)=<0.7,0.3>** **sampled: i^0**	**r=0.042** **$P(S\|i^0)$=<0.95,0.05>** **sampled: s^0**	**r=0.748** **$P(G\|d^1,i^0)$=<0.05,0.25,0.7>** **sampled: g^3**	**r=0.549** **$P(L\|g^3)$=<0.99,0.01>** **sampled: l^0**	**No**
r=0.49 **P(D)=<0.6,0.4>** **sampled: d^0**	**r=0.023** **P(I)=<0.7,0.3>** **sampled: i^0**	**r=0.849** **$P(S\|i^0)$=<0.95,0.05>** **sampled: s^0**	**r=0.009** **$P(G\|d^0,i^0)$=<0.3,0.4,0.3>** **sampled: g^1**		**Yes**
r=0.612 **P(D)=<0.6,0.4>** **sampled: d^1**	**r=0.936** **P(I)=<0.7,0.3>** **sampled: i^1**	**r=0.417** **$P(S\|i^1)$=<0.2,0.8>** **sampled: s^1**			**Yes**
r=0.63	**r=0**	**r=0.873** **$P(S\|i^0)$=<0.95,0.05>**	**r=0.453**	**r=0.89**	**No**

P(D)=<0.6,0.4> sampled: d^1	P(I)=<0.7,0.3> sampled: i^0	sampled: s^0	$P(G \mid d^1,i^0)$=<0.05,0.25,0.7> sampled: g^3	$P(L \mid g^3)$=<0.99,0.01> sampled: l^0	
r=0.768 P(D)=<0.6,0.4> sampled: d^1	r=0.859 P(I)=<0.7,0.3> sampled: i^1	r=0.394 $P(S \mid i^1)$=<0.2,0.8> sampled: s^1			Yes
r=0.157 P(D)=<0.6,0.4> sampled: d^0	r=0.103 P(I)=<0.7,0.3> sampled: i^0	r=0.81 $P(S \mid i^0)$=<0.95,0.05> sampled: s^0	r=0.139 $P(G \mid d^0,i^0)$=<0.3,0.4,0.3> sampled: g^1		Yes
r=0.037 P(D)=<0.6,0.4> sampled: d^0	r=0.12 P(I)=<0.7,0.3> sampled: i^0	r=0.51 $P(S \mid i^0)$=<0.95,0.05> sampled: s^0			Neither Yes nor No

b. What is $P(I|S=s^0, G=g^3)$ with respect to your sample?

I	Count	Normalize
i^0	2	1
i^1	0	0

$P(I|S=s^0, G=g^3)$ = <1, 0>

3. Likelihood Weighting: Assume the same evidence as in question 2. That is, $S=s^0$, $G=g^3$. Perform likelihood weighting.

a. **Complete the following table. Continue sampling until you run out of random numbers. Add rows as needed. The first row is given as an example.**

D	I	$S=s^0$	$G=g^3$	L	Weight
r=0.489 $P(D)=<0.6,0.4>$ sampled: d^0	r=0.833 $P(I)=<0.7,0.3>$ sampled: i^1	No sampling. $w=1*P(s^0\|i^1)$ w=1*0.2=0.2	No sampling. $w=0.2*P(g^3\|d^0,i^1)$ w=0.2*0.02=0.004	r=0.869 $P(L\|g3)=<0.99,0.01>$ sampled: l^0	0.004
r=0.652 $P(D)=<0.6,0.4>$ sampled: d^1	r=0.581 $P(I)=<0.7,0.3>$ sampled: i^0	No sampling. $w=1*P(s^0\|i^0)$ w=1*0.95 = 0.95	No sampling. $w=0.95*P(g^3\|d^1,i^0)$ w=0.95*0.7=0.665	r = 0.042 $P(L\|g3)=<0.99,0.01>$ sampled: l^0	0.665
r=0.748 $P(D)=<0.6,0.4>$ sampled: d^1	r = 0.549 $P(I)=<0.7,0.3>$ Sampled: i^0	No sampling. $w=1*P(s^0\|i^0)$ w=1*0.95 = 0.95	No sampling. $w=0.95*P(g^3\|d^1,i^0)$ w=0.95*0.7=0.665	r = 0.49 $P(L\|g3)=<0.99,0.01>$ sampled: l^0	0.665
r=0.023 $P(D)=<0.6,0.4>$ sampled: d^0	r = 0.849 $P(I)=<0.7,0.3>$ Sampled: i^1	No sampling. $w=1*P(s^0\|i^1)$ w=1*0.2 = 0.2	No sampling. $w=0.2*P(g^3\|d^0,i^1)$ w=0.2*0.02=0.004	r = 0.009 $P(L\|g3)=<0.99,0.01>$ sampled: l^0	0.004
r=0.612 $P(D)=<0.6,0.4>$ sampled: d^1	r = 0.936 $P(I)=<0.7,0.3>$ Sampled: i^1	No sampling. $w=1*P(s^0\|i^1)$ w=1*0.2 = 0.2	No sampling. $w=0.2*P(g^3\|d^1,i^1)$ w=0.2*0.2=0.04	r = 0.417 $P(L\|g3)=<0.99,0.01>$ sampled: l^0	0.04
r=0.63 $P(D)=<0.6,0.4>$ sampled: d^1	r = 0 $P(I)=<0.7,0.3>$ Sampled: i^0	No sampling. $w=1*P(s^0\|i^0)$ w=1*0.95 = 0.95	No sampling. $w=0.95*P(g^3\|d^1,i^0)$ w=0.95*0.7=0.665	r = 0.873 $P(L\|g3)=<0.99,0.01>$ sampled: l^0	0.665
r=0.453 $P(D)=<0.6,0.4>$ sampled: d^0	r = 0.89 $P(I)=<0.7,0.3>$ Sampled: i^1	No sampling. $w=1*P(s^0\|i^1)$ w=1*0.2 = 0.2	No sampling. $w=0.95*P(g^3\|d^0,i^1)$ w=0.2*0.02=0.004	r = 0.768 $P(L\|g3)=<0.99,0.01>$ sampled: l^0	0.004
r=0.859	r = 0.394	No sampling.	No sampling.	r = 0.157	0.665

P(D)=<0.6,0.4> sampled: d^1	P(I)=<0.7,0.3> Sampled: i^0	$w=1*P(s^0\|i^0)$ w=1*0.95 = 0.95	$w=0.95*P(g^3\|d^1,i^0)$ w=0.95*0.7=0.665	P(L|g3)=<0.99,0.01> sampled: l^0	
r=0.103 P(D)=<0.6,0.4> sampled: d^0	r = 0.81 P(I)=<0.7,0.3> Sampled: i^1	No sampling. $w=1*P(s^0\|i^1)$ w=1*0.2 = 0.2	No sampling. $w=0.2*P(g^3\|d^0,i^1)$ w=0.2*0.02=0.004	r = 0.139 P(L|g3)=<0.99,0.01> sampled: l^0	0.004
r=0.037 P(D)=<0.6,0.4> sampled: d^0	r = 0.12 P(I)=<0.7,0.3> Sampled: i^0	No sampling. $w=1*P(s^0\|i^0)$ w=1*0.95 = 0.95	No sampling. $w=0.95*P(g^3\|d^0,i^0)$ w=0.95*0.3=0.285	r = 0.51 P(L|g3)=<0.99,0.01> sampled: l^0	0.285

b. What is $P(I|S=s^0, G=g^3)$ with respect to your sample?

I	Count	Weight	Count*Weight
i^0	4	0.665	2.66
i^0	1	0.285	0.285
i^1	4	0.004	0.016
i^1	1	0.04	0.04

I	Sum	Normalize
i^0	2.945	2.945/3.001 = 0.98
i^1	0.056	0.056/3.001 = 0.02

$P(I|S=s^0, G=g^3)$ = <0.98, 0.02>

4. Gibbs Sampling for Bayesian Networks – with no evidence: Given no evidence, perform Gibbs sampling on the student network, using the random numbers in the random.csv file. Assume we initialize the network as follows: $D=d^0$, $I=i^0$, $S=s^0$, $G=g^2$, $L=l^1$. Visit the variables in the following order: D, I, S, G, L. Complete the following table. Continue sampling until you complete the first three iterations only. The first row is the initialization. Feel free to reuse probability computations as appropriate.

	D	**I**	**S**	**G**	**L**
Initialization	d^0	i^0	s^0	g^2	l^1
Iteration 1	**r = 0.489** $P(D\|i^0, s^0, g^2, l^1)$ $\propto P(D)P(g^2\|D,i^0)$ => <(0.6*0.4), (0.4*0.25)> => <0.24, 0.1> **Normalize** **<0.71, 0.29>** **sampled: d^0**	**r=0.833** $P(I\|d^0,s^0, g^2, l^1) \propto$ $P(I)P(s^0\|I)P(g^2\|d^0,I)$ =><(0.7*0.95*0.4), (0.3*0.2*0.08)> =><0.266, 0.0048> **Normalize** **<0.98, 0.02>** **sampled: i^0**	**r=0.869** $P(S\|d^0,i^0, g^2, l^1) \propto$ $P(S\|i^0)$ =><(0.95), (0.05)> **Normalize** **<0.95, 0.05>** **sampled: s^0**	**r=0.652** $p(G\|d^0,i^0,s^0, l^1) \propto$ $P(G\|d^0,i^0)P(l^1\|G)$ =><(0.3*0.9), (0.4*0.6), (0.3*0.01)> =><0.27, 0.24, 0.003> **Normalize** **<0.526, 0.467, 0.005>** **sampled: g^2**	**r=0.581** $P(L\|d^0,i^0,s^0,g^2)$ $\propto P(L\|g^2)$ =><(0.4), (0.6)> **Normalize** **<0.4, 0.6>** **sampled: l^1**
Iteration 2	**r=0.042** $P(D\|i^0, s^0, g^2, l^1)$ $\propto P(D)P(g^2\|D,i^0)$ => <(0.6*0.4), (0.4*0.25)> => <0.24, 1> **Normalize** **<0.71, 0.29>** **sampled: d^0**	**r=0.748** $P(I\|d^0,s^0, g^2, l^1) \propto$ $P(I)P(s^0\|I)P(g^2\|d^0,I)$ =><(0.7*0.95*0.4), (0.3*0.2*0.08)> =><0.266, 0.0048> **Normalize** **<0.98, 0.02>** **sampled: i^0**	**r=0.549** $P(S\|d^0,i^0, g^2, l^1) \propto$ $P(S\|i^0)$ =><(0.95), (0.05)> **Normalize** **<0.95, 0.05>** **sampled: s^0**	**r=0.49** $p(G\|d^0,i^0,s^0, l^1) \propto$ $P(G\|d^0,i^0)P(l^1\|G)$ =><(0.3*0.9), (0.4*0.6), (0.3*0.01)> =><0.27, 0.24, 0.003> **Normalize** **<0.526, 0.467, 0.005>** **sampled: g^1**	**r=0.023** $P(L\|d^0,i^0,s^0,g^1)$ $\propto P(L\|g^1)$ =><(0.1), (0.9)> **Normalize** **<0.1, 0.9>** **sampled: l^0**
Iteration 3	**r= 0.849** $P(D\|i^0, s^0, g^1, l^0)$ $\propto P(D)P(g^1\|D,i^0)$ =><(0.6*0.3), (0.4*0.05)> =><0.18, 0.02> **Normalize** **<0.9, 0.1>** **sampled: d^0**	**r= 0.009** $P(I\|d^0,s^0, g^1, l^0) \propto$ $P(I)P(s^0\|I)P(g^1\|d^0,I)$ =><(0.7*0.95*0.3), (0.3*0.2*0.9)> =><0.1995, 0.054> **Normalize** **<0.79, 0.21>** **sampled: i^0**	**r=0.612** $P(S\|d^0,i^0, g^1, l^0) \propto$ $P(S\|i^0)$ =><(0.95), (0.05)> **Normalize** **<0.95, 0.05>** **sampled: s^0**	**r=0.936** $p(G\|d^0,i^0,s^0, l^0) \propto$ $P(G\|d^0,i^0)P(l^0\|G)$ =><(0.3*0.1), (0.4*0.4), (0.3*0.99)> =><0.03, 0.16, 0.297> **Normalize** **<0.06, 0.33, 0.61>** **sampled: g^3**	**r=0.417** $P(L\|d^0,i^0,s^0,g^3)$ $\propto P(L\|g^3)$ =><0.99, 0.01> **Normalize** **<0.99, 0.01>** **sampled: l^0**

5. Gibbs Sampling for Bayesian Networks – with evidence: Assume the evidence is $S=s^0$, $G=g^3$. Perform Gibbs sampling on the student network, using the random numbers in the random.csv file. Assume we initialize the remaining variables as follows $D=d^0$, $I=i^0$, $L=l^1$. Visit the variables in the following order: D, I, L. Complete the following table. Continue sampling until you complete the first three iterations only. The first row is the initialization. Feel free to reuse probability computations as appropriate.

	D	**I**	**$S=s^0$**	**$G=g^3$**	**L**
Initialization	d^0	i^0	s^0	g^3	l^1
Iteration 1	**r = 0.489** $P(D\|i^0, s^0, g^3, l^1) \propto$ $P(D)P(g^3\|D,i^0)$ => <(0.6*0.3), (0.4*0.7)> => <0.18, 0.28> **Normalize** **<0.39, 0.61>** **sampled: d^1**	**r=0.833** $P(I\|d^1,s^0, g^3, l^1) \propto$ $P(I)P(s^0\|I)P(g^3\|d^1,I)$ =><(0.7*0.95*0.7), (0.3*0.2*0.2)> =><0.4655, 0.012> **Normalize** **<0.97, 0.03>** **sampled: i^0**	Not sampled	Not sampled	**r=0.869** $P(L\|d^1,i^0,s^0,g^3) \propto P(L\|g^3)$ =><(0.99), (0.01)> **Normalize** **=><0.99, 0.01>** **sampled: l^0**
Iteration 2	**r=0.652** $P(D\|i^0, s^0, g^3, l^1) \propto$ $P(D)P(g^3\|D,i^0)$ => <(0.6*0.3), (0.4*0.7)> => <0.18, 0.28> **Normalize** **<0.39, 0.61>** **sampled: d^1**	**r=0.581** $P(I\|d^1, s^0, g^3, l^1) \propto$ $P(I)P(s^0\|I)P(g^3\|d^1,I)$ =><(0.7*0.95*0.7), (0.3*0.2*0.2)> =><0.4655, 0.012> **Normalize** **<0.97, 0.03>** **sampled: i^0**	Not sampled	Not sampled	**r=0.042** $P(L\|d^1,i^0,s^0,g^3) \propto P(L\|g^3)$ =><(0.99), (0.01)> **Normalize** **=><0.99, 0.01>** **sampled: l^0**
Iteration 3	**r= 0.748** $P(D\|i^0, s^0, g^3, l^0) \propto$ $P(D)P(g^3\|D,i^0)$ =><(0.6*0.3), (0.4*0.7)> => <0.18, 0.28> **Normalize** **<0.39, 0.61>** **sampled: d^1**	**r= 0.549** $P(I\|d^1,s^0, g^3, l^0) \propto$ $P(I)P(s^0\|I)P(g^3\|d^1,I)$ =><(0.7*0.95*0.7), (0.3*0.2*0.2)> =><0.4655, 0.012> **Normalize** **<0.97, 0.03>** **sampled: i^0**	Not sampled	Not sampled	**r=0.49** $P(L\|d^1,i^0,s^0,g^3) \propto P(L\|g^3)$ =><(0.99), (0.01)> **Normalize** **=><0.99, 0.01>** **sampled: l^0**

6. Gibbs Sampling for Markov Networks – with no evidence: For this and the following question, assume a fully connected Markov network over variables A, B, and C. Assume the following factors. Perform Gibbs sampling on this network, using the random numbers in random.csv. Assume we initialize the variables as follows: A=T, B=F, C=F. Visit the variables in the following order: A, B, C.

A	ϕ(A)
T	2
F	1

B	ϕ(B)
T	1
F	4

C	ϕ(C)
T	1
F	8

A	B	ϕ(A,B)
T	T	5
T	F	1
F	T	1
F	F	5

A	C	ϕ(A,C)
T	T	6
T	F	1
F	T	1
F	F	6

B	C	ϕ(B,C)
T	T	1
T	F	10
F	T	10
F	F	1

Complete the following table. Continue sampling until you complete the first three iterations only. The first row is the initialization. Feel free to reuse probability computations as appropriate.

	A	**B**	**C**
Initialization	T	F	F
Iteration 1	r=0.489 $P(A \mid B=f, C=f) \propto \phi(A) * \phi(A, B=f) * \phi(A, C=f)$ => <(2*1*1), (1*5*6)> => <2, 30> **Normalize:** **$P(A \mid B=f, C=f)$ = <0.06, 0.94>** **sampled: A=f**	r=0.833 $P(B \mid A=f, C=f) \propto \phi(B) * \phi(A=f, B) * \phi(B, C=f)$ => <(1*1*10), (4*5*1)> => <10, 20> **Normalize:** **$P(B \mid A=f, C=f)$ = <0.33, 0.66>** **sampled: B=f**	r=0.869 $P(C \mid A=f, B=f) \propto \phi(C) * \phi(A=f, C) * \phi(B=f, C)$ => <(1*1*10), (8*6*1)> => <10, 48> **Normalize:** **$P(C \mid A=f, B=f)$ = <0.17, 0.83>** **Sampled: C=f**
Iteration 2	r=0.652 $P(A \mid B=f, C=f) \propto \phi(A) * \phi(A, B=f) * \phi(A, C=f)$ => <(2*1*1), (1*5*6)> => <2, 30> **Normalize**: **$P(A \mid B=f, C=f)$ = <0.06, 0.94>** **sampled: A=f**	r=0.581 $P(B \mid A=f, C=f) \propto \phi(B) * \phi(A=f, B) * \phi(B, C=f)$ => <(1*1*10), (4*5*1)> => <10, 20> **Normalize**: **$P(B \mid A=f, C=f)$ = <0.33, 0.66>** **sampled: B=f**	r=0.042 $P(C \mid A=f, B=f) \propto \phi(C) * \phi(A=f, C) * \phi(B=f, C)$ => <(1*1*10), (8*6*1)> => <10, 48> **Normalize**: **$P(C \mid A=f, B=f)$ = <0.17, 0.83>** **sampled: C=t**
Iteration 3	r=0.748 $P(A \mid B=f, C=t) \propto \phi(A) * \phi(A, B=f) * \phi(A, C=t)$ => <(2*1*6), (1*5*1)> => <12, 5> **Normalize**: **$P(A \mid B=f, C=t)$ = <0.71, 0.29>** **sampled: A=f**	r=0.549 $P(B \mid A=f, C=t) \propto \phi(B) * \phi(A=f, B) * \phi(B, C=t)$ => <(1*1*1), (4*5*10)> => <1, 200> **Normalize**: **$P(B \mid A=f, C=t)$ = <0.005, 0.995>** **sampled: B=f**	r=0.49 $P(C \mid A=f, B=f) \propto \phi(C) * \phi(A=f, C) * \phi(B=f, C)$ => <(1*1*10), (8*6*1)> => <10, 48> **Normalize:** **$P(C \mid A=f, B=f)$ = <0.17, 0.83>** **sampled: C=f**

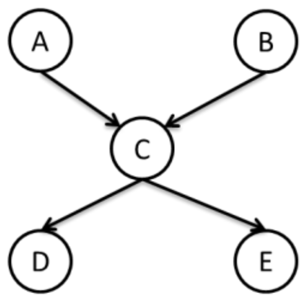
7. Gibbs Sampling for Markov Networks – with evidence: For this question, use above Markov network. Assume the evidence is B=T. Perform Gibbs sampling on this network, using the random numbers in random.csv. Assume we initialize the remaining variables as follows: A=T, C=F. Visit the variables in the following order: A, C. Complete the following table. Continue sampling until you complete the first three iterations only. The first row is the initialization. Feel free to reuse probability computations as appropriate.

	A	**B=T**	**C**
Initialization	T	T	F
Iteration 1	r=0.489 P(A|B=t, C=f) ∝ ϕ(A) * ϕ(A, B=t) * ϕ(A, C=f) => <(2*5*1), (1*1*6)> => <10, 6> **Normalize:** **P(A|B=t, C=f) = <0.625, 0.375>** **sampled: A= t**	Not sampled	r=0.833 P(C|A=t, B=t) ∝ ϕ(C) * ϕ(A=t, C) * ϕ(B=t, C) => <(1*6*1), (8*1*10)> => <6, 80> **Normalize:** **P(C|A=t, B=t) = <0.07, 0.93>** **sampled: C=f**
Iteration 2	r=0.869 P(A|B=t, C=f) ∝ ϕ(A) * ϕ(A, B=t) * ϕ(A, C=f) => <(2*5*1), (1*1*6)> => <10, 6> **Normalize:** **P(A|B=t, C=f) = <0.625, 0.375>** **sampled: A=f**	Not sampled	r=0.652 P(C|A=f, B=t) ∝ ϕ(C) * ϕ(A=f, C) * ϕ(B=t, C) => <(1*1*1), (8*6*10)> => <1, 480> **Normalize:** **P(C|A=f, B=t) = <0.002, 0.998>** **sampled: C=f**
Iteration 3	r=0.581 P(A|B=t, C=f) ∝ ϕ(A) * ϕ(A, B=t) * ϕ(A, C=f) => <(2*5*1), (1*1*6)> => <10, 6> **Normalize:** **P(A|B=t, C=f) = <0.625, 0.375>** **sampled: A=t**	Not sampled	r=0.042 P(C|A=t, B=t) ∝ ϕ(C) * ϕ(A=t, C) * ϕ(B=t, C) => <(1*6*1), (8*1*10)> => <6, 80> **Normalize:** **P(C|A=t, B=t) = <0.07, 0.93>** **sampled: C=t**

8. MAP for Bayesian networks: we are given the following Bayesian network. Answer the following questions.

A	P(A)
T	0.7
F	0.3

B	P(B)
T	0.2
F	0.8



A	B	P(C|A,B)
T	T	<0.4, 0.6>
T	F	<0.7, 0.3>
F	T	<0.5, 0.5>
F	F	<0.2, 0.8>

C	P(D|C)
T	<0.7, 0.3>
F	<0.2, 0.8>

C	P(E|C)
T	<0.1, 0.9>
F	<0.6, 0.4>

a. What is the MAP assignment, given no evidence?

Bayesian distribution for the above network is: P(A)P(B)P(C|A,B)P(D|C)P(E|C)

ELIMINATE E:

C	E	$\psi(C, E)$ = P(E|C)
T	T	0.1
T	F	0.9
F	T	0.6
F	F	0.4

τ_1(C) = <0.9, 0.6>

ELIMINATE D:

C	D	$\psi(C, D)$ = P(D|C)
T	T	0.7

T	F	0.3
F	T	0.2
F	F	0.8

τ_2(C) = <0.7, 0.8>

ELIMINATE C:

A	**B**	**C**	ψ(A, B, C) = P(C|A,B)τ_1(C)τ_2(C)
T	T	T	0.4 * 0.9 * 0.7 = 0.252
T	T	F	0.6 * 0.6 * 0.8 = 0.288
T	F	T	0.7 * 0.9 * 0.7 = 0.441
T	F	F	0.3 *0.6 * 0.8 = 0.144
F	T	T	0.5 * 0.9 * 0.7 = 0.315
F	T	F	0.5 * 0.6 * 0.8 = 0.24
F	F	T	0.2 * 0.9 * 0.7 = 0.126
F	F	F	0.8 * 0.6 * 0.8 = 0.384

A	B	τ_3(A, B)
T	T	0.288
T	F	0.441
F	T	0.315
F	F	0.384

ELIMINATE B:

A	B	$\psi(A, B) = P(B) * \tau_3(A, B)$
T	T	0.2 * 0.288 = 0.0576
T	F	0.8 * 0.441 = 0.3528
F	T	0.2 * 0.315 = 0.063
F	F	0.8 * 0.384 = 0.3072

τ_4(A) = <0.3528, 0.3072>

ELIMINATE A:

A	$\psi(A) = P(A) * \tau_4(A)$
T	0.7 * 0.3528 = 0.24696
F	0.3 * 0.3072 = 0.09216

MAP = A=T, B=F, C=T, D=T, E=F

b. What is the MAP assignment, given C=T?

Bayesian distribution for the above network is: P(A)P(B)P(C=t|A,B)P(D|C=t)P(E|C=t)

ELIMINATE E: τ_1() = <0.1, 0.9>

ELIMINATE D: τ_2() = <0.7, 0.3>

ELIMINATE B:

A	B	$\psi(A, B, C{=}t) = P(B) * P(C{=}t \mid A,B)$
T	T	0.2 * 0.4 = 0.08
T	F	0.8 * 0.7 = 0.56
F	T	0.2 * 0.5 = 0.01
F	F	0.8 * 0.2 = 0.16

$\tau_3(A)$ = <0.56, 0.16>

ELIMINATE A:

A	$\psi(A) = P(A) * \tau_3(A)$
T	0.7 * 0.56 = 0.392
F	0.3 * 0.16 = 0.048

MAP|C=t = A=t, B=f, D=t, E=f

c. **What is the MAP assignment to (A, B), given no evidence?**

MAP(A, B) = $\text{argmax}_{A, B} \sum_{C, D, E}$ P(A) * P(B) * P(C|A,B) * P(D|C) * P(E|C)

ELIMINATE E:

C	E	$\psi(C, E)$ = P(E|C)
T	T	0.1
T	F	0.9
F	T	0.6
F	F	0.4

$\tau_1(C)$ = <1, 1>

ELIMINATE D:

C	D	$\psi(C, D)$ = P(D|C)
T	T	0.7
T	F	0.3

F	T	0.2
F	F	0.8

$\tau_2(C)$ = <1, 1>

ELIMINATE C:

A	**B**	**C**	$\psi(A, B, C) =$ $P(C\|A,B)\tau_1(C)\tau_2(C)$
T	T	T	0.4 * 1 *1 = 0.4
T	T	F	0.6 * 1 * 1 = 0.6
T	F	T	0.7 * 1 * 1 = 0.7
T	F	F	0.3 * 1 * 1 = 0.3
F	T	T	0.5 * 1 * 1 = 0.5
F	T	F	0.5 * 1 * 1 = 0.5
F	F	T	0.2 * 1 * 1 = 0.2
F	F	F	0.8 * 1 * 1 = 0.8

A	B	$\tau_3(A, B)$
T	T	1
T	F	1
F	T	1
F	F	1

Final MAX product would be: P(A) P(B) τ_3(A, B)

A	B	P(A) * P(B) * τ_3(A, B)
T	T	0.7 * 0.2 * 1 = 0.14
T	**F**	**0.7 * 0.8 * 1 = 0.56**
F	T	0.3 * 0.2 * 1 = 0.06
F	F	0.3 * 0.8 * 1 = 0.24

MAP(A, B) = A=T, B=F

d. **What is the MAP assignment to (A, B), given D=T, E=F?**

MAP(A, B) = argmax$_{A,\,B}$ Σ_CP(A) * P(B) * P(C|A,B) * P(D=t|C) * P(E=f|C)

Bayesian distribution for the above network is: P(A)P(B)P(C|A,B)P(D=t|C)P(E=f|C)

ELIMINATE C:

ψ(A, B, C) = P(C|A,B)P(D=t|C)P(E=f|C)

A	**B**	**C**	**ψ(A, B, C) = P(C|A,B)P(D=t|C)P(E=f|C)**
T	T	T	0.4 * 0.7 * 0.9 = 0.252
T	T	F	0.6 * 0.2 * 0.4 = 0.048
T	F	T	0.7 * 0.7 * 0.9 = 0.441
T	F	F	0.3 * 0.2 * 0.4 = 0.024
F	T	T	0.5 * 0.7 * 0.9 = 0.315
F	T	F	0.5 * 0.2 * 0.4 = 0.04
F	F	T	0.2 * 0.7 * 0.9 = 0.126
F	F	F	0.8 * 0.2 * 0.4 = 0.064

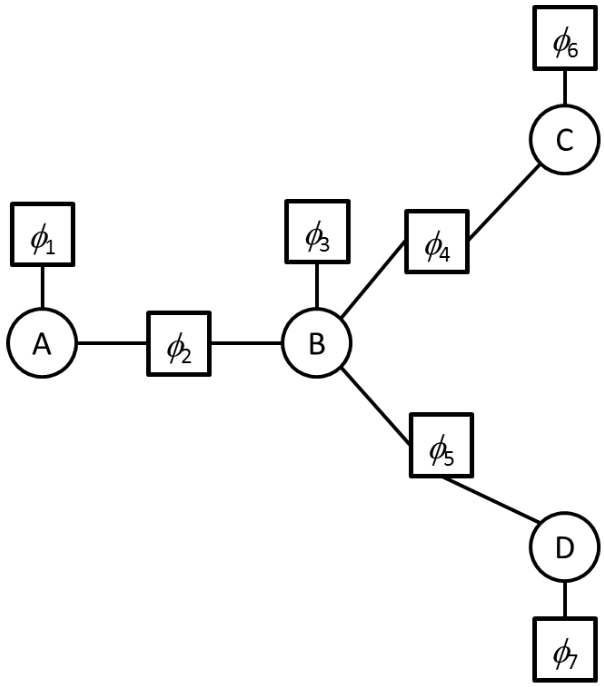
A	B	$\tau_1(A, B)$
T	T	0.3
T	F	0.465
F	T	0.355
F	F	0.19

$\psi(A, B) = P(A) * P(B) * \tau_1(A, B)$

A	B	$\psi(A, B) = P(A) * P(B) * \tau_1(A, B)$
T	T	0.7 * 0.2 * 0.3 = 0.042
T	**F**	**0.7 * 0.8 * 0.465 = 0.26**
F	T	0.3 * 0.2 * 0.355 = 0.0213
F	F	0.3 * 0.8 * 0.19 = 0.0456

MAP (A, B) | D=T, E=F = A=T, B=F

9. MAP for Markov networks: We are given the following factor graph and the potential functions. Answer the following questions. Do not use calculators.



f(condition) = 1 if condition is True, 0 otherwise

$\phi_1 = e^{f(A=T)-f(A=F)}$
$\phi_2 = e^{-f(A=T,\ B=T)+f(A=T,\ B=F)+f(A=F,\ B=T)-f(A=F,\ B=F)}$
$\phi_3 = e^{f(B=T)-f(B=F)}$
$\phi_4 = e^{f(B=T,\ C=T)-2f(B=T,\ C=F)-2f(B=F,\ C=T)+f(B=F,\ C=F)}$
$\phi_5 = e^{-3f(B=T,\ D=T)+f(B=T,\ D=F)+f(B=F,\ D=T)-4f(B=F,\ D=F)}$
$\phi_6 = e^{-2f(C=T)+f(C=F)}$
$\phi_7 = e^{f(D=T)-4f(D=F)}$

e. What is the MAP assignment, given no evidence?

Probability distribution for the above network is

P(A)*P(B)*P(C)*P(D)*P(A, B)*P(B, C)*P(C, D) = $\Phi_1 * \Phi_2 * \Phi_3 * \Phi_4 * \Phi_5 * \Phi_6 * \Phi_7$

ELIMINATE A:

A	B	ψ**(A, B) =** $\Phi_1 * \Phi_2 = e^{f(A=T)-f(A=F)} * e^{-f(A=T,\ B=T)+f(A=T,\ B=F)+F(A=F,\ B=T)-f(A+F,B=F)}$
T	T	$e^{1-0} * e^{-1-0+0-0} = e^1 * e^{-1} = e^0 = 1$
T	F	$e^{1-0} * e^{-0-1+0-0} = e^1 * e^1 = e^2$
F	T	$e^{0-1} * e^{-0-0+1-0} = e^{-1} * e^1 = e^0 = 1$
F	F	$e^{0-1} * e^{-0-0+0-1} = e^{-1} * e^{-1} = e^{-2}$

τ_1**(B) = <1,** e^2**>**

ELIMINATE C:

B	C	$\psi(B, C) = \Phi_6 * \Phi_4 = e^{-2f(C=T)+f(C=F)} * e^{f(B=T,\ C=T)-2f(B=T,\ C=F)-2f(B=F,\ C=T)+f(B=F,C=F)}$
T	T	$e^{-2+0} * e^{1-2*0-2*0-0} = e^{-2} * e^{1} = e^{-1}$
T	F	$e^{1} * e^{-2} = e^{-1}$
F	T	$e^{-2} * e^{-2} = e^{-4}$
F	F	$e^{1} * e^{1} = e^{2}$

$\tau_2(B) = <e^{-1}, e^{2}>$

ELIMINATE D:

B	D	$\psi(B, D) = \Phi_7 * \Phi_5 = e^{f(D=T)-4f(D=F)} * e^{-3f(B=T,\ D=T)+f(B=T,\ D=F)+f(B=F,\ D=T)-4f(B=F,\ D=F)}$
T	T	$e^{1} * e^{-3*1} = e^{-2}$
T	F	$e^{-4*1} * e^{1} = e^{-3}$
F	T	$e^{1} * e^{1} = e^{2}$
F	F	$e^{-4*1} * e^{-4*1} = e^{-8}$

$\tau_3(B) = <e^{-2}, e^{2}>$ $\tau_2(B) = <e^{-1}, 1>$

ELIMINATE B:

B	$\psi(B, D) = \Phi_3 * \tau_1(B) * \tau_2(B) * \tau_3(B) = e^{f(B=T)-f(B=F)} * \tau_1(B) * \tau_2(B) * \tau_3(B)$
T	$e^{1} * 1 * e^{-1} * e^{-2} = e^{-2}$
F	$e^{-1} * e^{2} * e^{2} * e^{2} = e^{5}$

MAP ORDER: B=F, D=T, C=F, A=T

f. What is the MAP assignment, given C=T, D=T?

Probability distribution for the above network is

$P(A)*P(B)*P(C=T)*P(D=T) = \phi_1 * \phi_2 * \phi_3 * \phi_4 * \phi_5 * \phi_6 * \phi_7$

ELIMINATE A:

A	B	$\psi(A, B) = \Phi_1 * \Phi_2 = e^{f(A=T)-f(A=F)} * e^{-f(A=T, B=T)+f(A=T, B=F)+F(A=F, B=T)-f(A+F,B=F)}$
T	T	$e^{1-0} * e^{-1-0+0-0} = e^1 * e^{-1} = e^0 = 1$
T	F	$e^{1-0} * e^{-0-1+0-0} = e^1 * e^1 = e^2$
F	T	$e^{0-1} * e^{-0-0+1-0} = e^{-1} * e^1 = e^0 = 1$
F	F	$e^{0-1} * e^{-0-0+0-1} = e^{-1} * e^{-1} = e^{-2}$

$\tau_1(B) = <1, e^2>$

ELIMINATE B:

$\Phi_2 = e^{f(B=T)-f(B=F)}$

$\Phi_4 \mid (C=T) = e^{f(B=T, C=T)-2 * 0 - 2f(B=F, C=T)+0} = e^{f(B=T, C=T)-2f(B=F, C=T)}$

$\Phi_5 \mid (D=T) = e^{-3f(B=T, D=T)+0+f(B=F, D=T)-4 * 0} = e^{-3f(B=T, D=T)+f(B=F, D=T)}$

$\Phi_6 \mid (C=T) = e^{-2 * 1} = e^{-2}$

$\Phi_7 \mid (D=T) = e^{1-4 * 0} = e^1$

B	$\psi(B) = \Phi_3 * \tau_1(B) * \Phi_4 * \Phi_5 * \Phi_6 * \Phi_7$
T	$e^1 * 1 * e^1 * e^{-3} * e^{-2} * e^1 = e^{-2}$
F	$e^{-1} * e^2 * e^{-2} * e^1 * e^{-2} * e^1 = e^{-1}$

MAP | (C=T, D=T) = B=F, A=T