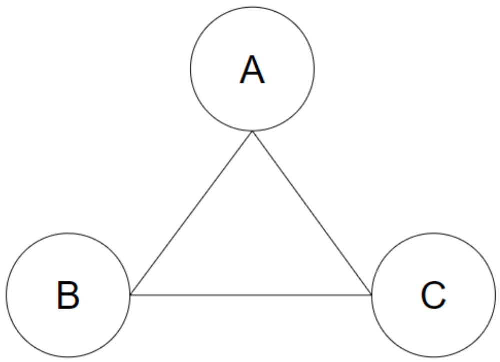
Tania Soutonglang

CS 583

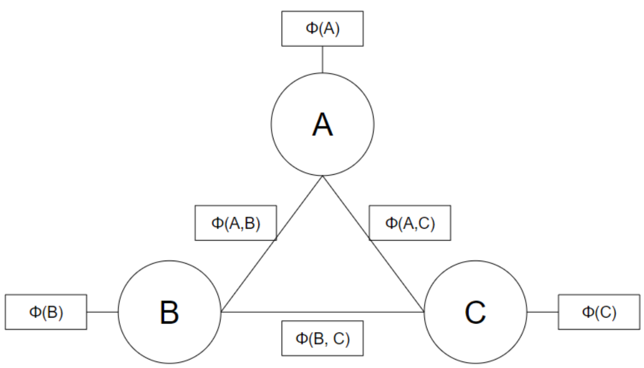
February 27, 2024

Assignment 2

1. We have three random variables, A, B, C, each of which is binary. We have the following factors over these variables: $\phi(A), \phi(B), \phi(C), \phi\ (A, B), \phi(A, C), \phi(B, C)$
   a. Draw a Markov network graph over these variables.



   b. Draw a factor graph over these variables.



   c. Here are the values of the factors. Compute P(A, B, C).

A	ϕ(A)
T	2
F	1

B	ϕ(B)
T	1
F	4

C	ϕ(C)
T	1
F	8

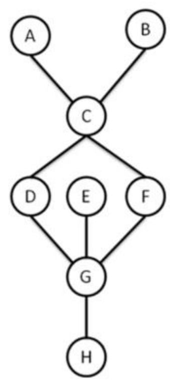
A	B	ϕ(A,B)
T	T	5
T	F	1
F	T	1
F	F	5

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

A	B	C	f(A)	f(B)	f(C)	f(A, B)	f(A, C)	f(B, C)	f(D) = f(A) f(B) f(C) f(A,B) f(A,C) f(B,C)	P(A, B, C) = 1/3045
T	T	T	2	1	1	5	6	1	60	0.01970
T	T	F	2	1	8	5	1	10	800	0.26273
T	F	T	2	4	1	1	6	10	480	0.15764
T	F	F	2	4	8	1	1	1	64	0.02102
F	T	T	1	1	1	1	1	1	1	0.00033
F	T	F	1	1	8	1	6	10	480	0.15764
F	F	T	1	4	1	5	1	10	200	0.06568
F	F	F	1	4	8	5	6	1	960	0.31527
									Z = 3045	

2. For the following Markov network graph, indicate whether the following independence statements are True or False.



   a. A ⊥ B FALSE
   b. A ⊥ B | C TRUE
   c. A ⊥ G | D FALSE
   d. A ⊥ G | D, F TRUE
   e. A ⊥ H | G TRUE

3. We have the Markov network over 3 binary variables: $A - B - C$. We define a pairwise Markov random field (MRF) over this network. We define the following features.

   $f_1(A) = 1\ if\ A = T$, 0 otherwise. $w_1 = \ln(2)$.

   $f_2(B) = 1\ if\ B = F$, 0 otherwise. $w_2 = -\ln(3)$.

   $f_3(C) = 1\ if\ C = T$, 0 otherwise. $w_3 = \ln(4)$.

   $f_4(A, B) = 1\ if\ A = B$, 0 otherwise. $w_4 = -\ln(5)$.

$f_5(B,C) = 1 \; if \; B \neq C$, 0 otherwise. $w_5 = \ln(6)$.

$$P(A,B,C) = \frac{1}{Z} e - \sum_{i=1}^{5} w_i \times f_i(A,B,C)$$

Note that to simplify the notation, we simply wrote $f_i(A,B,C)$, though features are defined only over nodes and edges. Assume f_i ignores the variables that it is not defined over. For example, $f_1(A,B,C) = f_1(A)$. Populate the following table.

Hints: $e^{\ln(x)} = x$ and $e^{-\ln(x)} = \frac{1}{e^{\ln(x)}} = \frac{1}{x}$.

A	B	C	f(A)	f(B)	f(C)	f(A,B)	f(B,C)	$e^{-\sum_{i=1}^{5} w_i \times f_i(A,B,C)}$	P(A, B, C)
T	T	T	0.69315	0.00000	1.38629	-1.60944	0	0.62500	0.03352
T	T	F	0.69315	0.00000	0	-1.60944	1.79176	0.41667	0.02235
T	F	T	0.69315	-1.09861	1.38629	0	1.79176	0.06250	0.00335
T	F	F	0.69315	-1.09861	0	0	0	1.50000	0.08045
F	T	T	0.00000	0.00000	1.38629	0	0	0.25000	0.01341
F	T	F	0.00000	0.00000	0	0	1.79176	0.16667	0.00894
F	F	T	0.00000	-1.09861	1.38629	-1.60944	1.79176	0.62500	0.03352
F	F	F	0.00000	-1.09861	0	-1.60944	0	15.00000	0.80447
								18.64583	

4. We have a document classification task. We have four documents, D1 and D2. We are interested in classifying these documents into one of two topics: Artificial Intelligence (AI) or Databases (DB). Each document has one or both words: Agent and SQL. We will represent this data as follows: the labels of D_1 and D_2 are: Y_1 and Y_2. Each Y_i can take one of two values: AI or DB. The presence of word Agent in document D_i represented as X_{1i}, and the presence of the word SQL in document D_i represented as X_{2i}. The documents and their contents are as follows:
   D_1: $X_{11} = True, X_{21} = False$ (i.e, D_1 contains only the word Agent)
   D_2: $X_{12} = True, X_{22} = True$ (i.e, D_2 contains both words)
   We also have the additional knowledge that D_1 cites D_2. We construct the following CRF, with the following feature functions.
   Features:

$f_1(X_{1i}, Y_i) = 1$ if $X_{1i} = T$ and $Y_i = AI$, 0 otherwise; $w_1 = -1$.

$f_2(X_{1i}, Y_i) = 1$ if $X_{1i} = T$ and $Y_i = DB$, 0 otherwise; $w_2 = +1$.

$f_3(X_{2i}, Y_i) = 1$ if $X_{2i} = T$ and $Y_i = AI$, 0 otherwise; $w_3 = +1$.

$f_4(X_{2i}, Y_i) = 1$ if $X_{2i} = T$ and $Y_i = DB$, 0 otherwise; $w_4 = -1$.

$f_5(Y_i, Y_j) = 1$ if $Y_i=Y_j$, 0 otherwise; $w_5 = -1$.

$f_6(Y_i, Y_j) = 1$ if $Y_i \neq Y_j$, 0 otherwise; $w_6 = +1$.

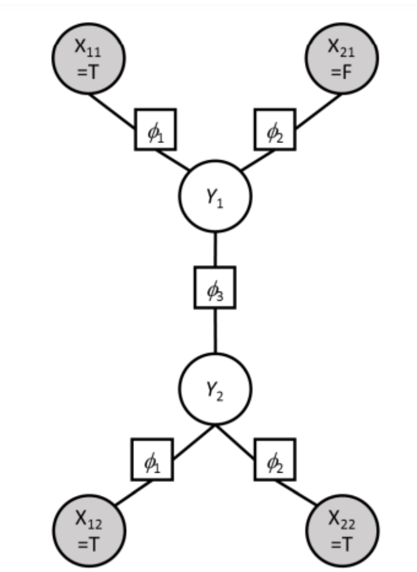
Potentials:

$\phi_1(X_{1i}, Y_i) = e^{-(w1f1+ w2f2)}$

$\phi_2(X_{2i}, Y_i) = e^{-(w3f3+ w4f4)}$

$\phi_3(Y_i, Y_j) = e^{-(w5f5+ w6f6)}$

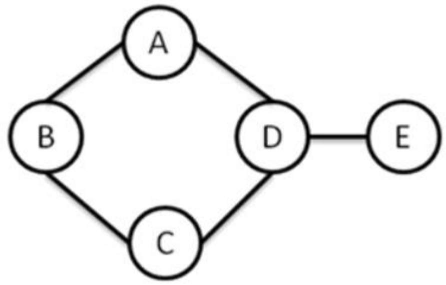
The factor graph is given as follows:



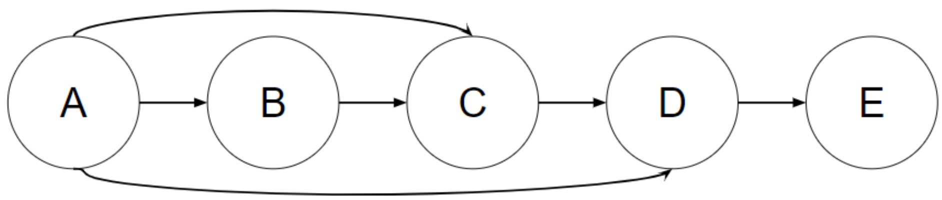
Calculate $P(Y|X)$, that is $P(Y_1, Y_2|X_{11} = T, X_{21} = F, X_{12} = T, X_{22} = T)$. What is the MAP assignment to Y given X? Show your work.

Y1	Y2	f(Y1, Y2)	f(Y1, X11=T)	f(Y1, X21=F)	f(Y2, X12=T)	f(Y2, X22=T)	f	P(Y|X)
AI	AI	2.71828	2.71828	1.00000	2.71828	0.36788	7.38906	0.77580
AI	DB	0.36788	2.71828	1.00000	0.36788	2.71828	1.00000	0.10499
DB	AI	0.36788	0.36788	1.00000	2.71828	0.36788	0.13534	0.01421
DB	DB	2.71828	0.36788	1.00000	0.36788	2.71828	1.00000	0.10499
							9.52439	

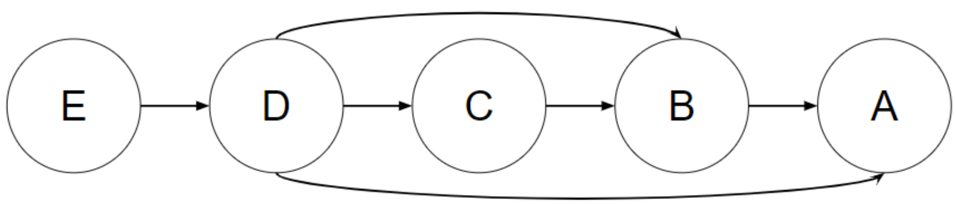
5. We are given the following Markov network structure, H.



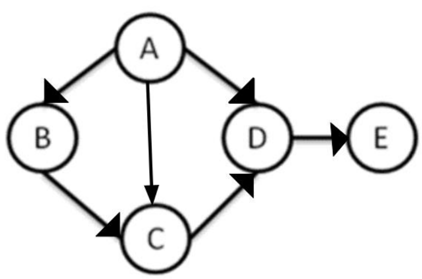
a. Find a minimal I-Map Bayesian network structure G_1 for H. Use the variable order of A, B, C, D, E.



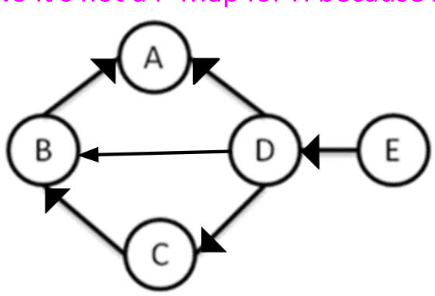
b. Find a minimal I-Map Bayesian network structure G_1 for H. Use the variable order of E, D, C, B, A.



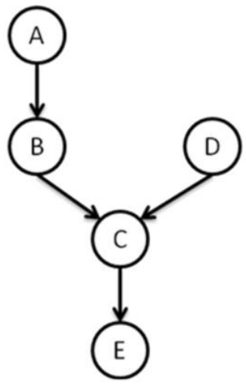
c. Is G_1 a P-Map for H? If not, which independencies are missing?
No it's not a P-Map for H



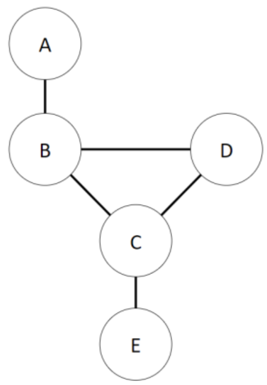
d. Is G_2 a P-Map for H? If not, which independencies are missing?
No it's not a P-Map for H because $B \perp D, E \mid A, C$ is not true.



6. We are given the following Bayesian network structure G.



a. Find a minimal I-Map Markov network structure H for G. Use any method/variable order you like; make sure H is a *minimal* I-Map.



b. Is H a P-Map for G? If not, which independencies are missing?
No, $A, B \perp D$ is not true.