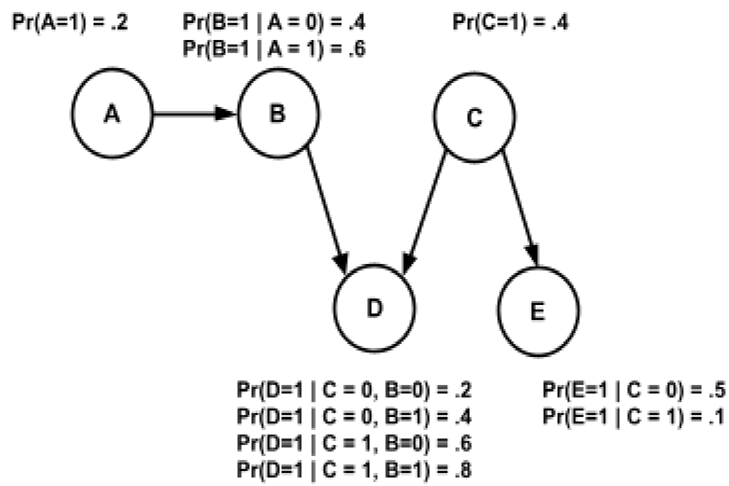
Homework 2

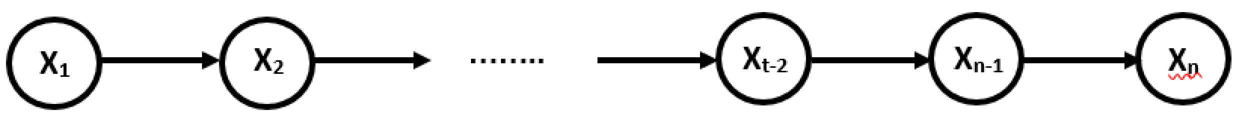
Q1- Bayesian Nets

Given the following network, calculate the probabilities below:

- P(E=1)
- P(C=1)
- P(C=1 | B=1)
- P(C=1 | B=0, D=0)
- P(D=0 | A=1)
- P(D=0 | B=0, E=1)



Q2- Markov chain



Consider a Markov Chain as above. Prove

A) $\mathbf{X_t} \perp \mathbf{X_{t-3}} \mid \mathbf{X_{t-1}}, \mathbf{X_{t-2}}$
B) $\mathbf{X_t} \perp \mathbf{X_{t-3}} \mid \mathbf{X_{t-2}}$
C) $\mathbf{X_t} \perp \mathbf{X_s} \mid \mathbf{X_{t-2}}$ for $\mathbf{s \leq t-3}$
D) $\mathbf{X_t} \perp \mathbf{X_s} \mid \mathbf{X_r}$ for $\mathbf{s<r<t}$
E) Given $\mathbf{X_{t-1}}$ and $\mathbf{X_{t+1}}$, $\mathbf{X_t}$ is conditionally independent of all other nodes.

You are not allowed to use the active trail or d-separation theorems (of course $\mathbf{X_{t-2}}$ separates $\mathbf{X_t}$ and $\mathbf{X_{t-3}}$). You can only make use of the following:

- Each node is independent of its non-descendants given its parents, and
- The joint distribution can be written as the product of the CPDs.

Q3- Markov Random Fields

Consider the following Markov Random Field over variables $A, B, C, D \in \{-1, 1\}$. The potential functions are

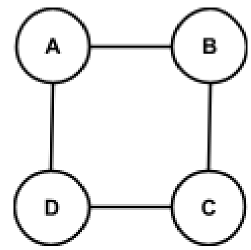
$\phi_1(A, B) = \exp(1(A = B))$

$\phi_2(B, C) = exp(-BC)$

$\phi_3(C, D) = exp(D - CD)$

$\phi_4(A, D) = exp(1(A \neq D))$

where, $1(\cdot)$ is the indicator function.



$P(A, B, C, D) = 1/Z\ \phi_1(A, B)\ \phi_2(B, C)\ \phi_3(C, D)\ \phi_4(A, D)$

1. obtain:
   a. the unnormalized measure $\tilde{P}(A, B, C) = Z\, P(A, B, C)$
   b. the unnormalized measure $\tilde{P}(A, B) = Z\, P(A, B)$
   c. the unnormalized measure $\tilde{P}(A) = Z\, P(A)$
   d. the partition function Z (using the fact that $\sum_{A=-1}^{1} P(A) = 1$.)
   - **in each case, simply your solution as much as you can**

2. Having Z, obtain the (normalized) distributions $P(A, B, C)$, $P(A, B)$,$P(A)$
3. Derive $P(A \mid B, C, D)$, and $P(A \mid B, D)$. Show that **A** is independent of **C** given **B,D**.

The Hammersley-Clifford theorem

An undirected graphical model with a set of nodes G and the neighbourhood system $N$is called a Markov Random Field if

$$p(X_i \mid X_{G\setminus\{i\}}) = p(X_i \mid X_{N_i}) \qquad (1)$$

where G is the set of nodes of the graph, $G \setminus \{i\}$ represents all graph nodes except node i, and N_i denotes the neighbours of node i.

An undirected graphical model is called a Gibbs Random Field (and its joint distribution a Gibbs distribution) if the corresponding joint distribution can be factorized as the product of functions over cliques (=fully connected subgraphs) of the graph

$$p(X_G) = \frac{1}{Z} \prod_{c \in C} \Phi_c(X_c) \qquad (2)$$

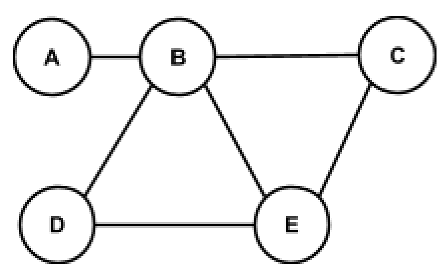
where C is the set of all cliques (or a subset of all cliques) and X_c is the set of variables in the clique c. The Hammersley-Clifford theorem states that if the joint distribution $p(X_G)$ is nonzero for all X_G, then an undirected graphical model is a Markov random field if and only if it is a Gibbs random field. In other words, the two models are equivalent.

Q5- Prove the easy direction of Hammersley-Clifford (Gibbs -> MRF)

Show that if the joint distribution $P(A, B, C, D, E)$ is **positive** and can be written as the product of factors over the graph cliques, as in Eq. (2) above, then each node in the graph is independent of the non-neighbouring nodes given its neighbours, as Eq. (1).

Hint: Dervie $p(X_i \mid X_{G\setminus\{i\}})$ and $p(X_i \mid X_{N_i})$ and show that they are equal.

Q6- Prove the hard direction of Hammersley-Clifford (MRF -> Gibbs) for a simple graph



Show that given the conditional independence relations implied by the following MRF graph, the corresponding joint distribution can be written as a product of factors over cliques (and hence is a Gibbs distribution). That is, there exits factors Φ_1, Φ_2, Φ_3 such that

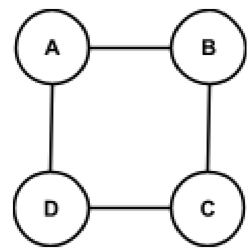
$$P(A, B, C, D, E) = \Phi_1(A, B)\, \Phi_2(B, E, D)\, \Phi_3(B, E, C)$$

Hint:

1- Start by writing $P(A, B, C, D, E) = P(A \mid B, C, D, E)\; P(B, C, D, E)$ and using the Makov property.

2- Show that if $P(X, Y, Z \mid T) = P(X \mid T)\ P(Y, Z \mid T)$ then we have $P(X, Y \mid T) = P(X \mid T)\ P(Y \mid T)$. Using this, prove that $D \perp C \mid B, E$ (notice that the Markov property gives $D \perp (C, A) \mid B, E$.

Not part of the homework!

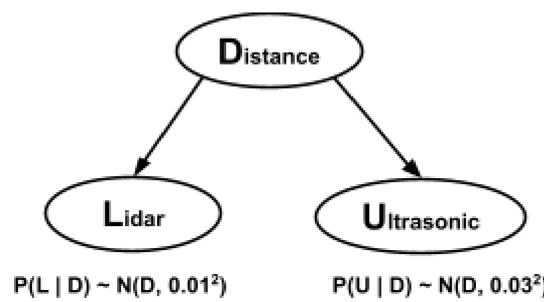


To see that the Hammersley-Clifford is not so trivial, try to solve question 6 for the following graph. See if you can show that the joint distribution can be factorized as

$$P(A, B, C, D) = \Phi_1(A, B)\ \Phi_2(B, c)\ \Phi_3(C, D)\ \Phi_4(D, A),$$

given only the Markov property.

Q7- Simple sensor fusion



Suppose we have obtained distance measurements using a LiDAR and an Ultrasound sensor. The LiDAR records a distance of 2.24 meters, while the ultrasound sensor gives 2.13 meters. We assume both sensors' errors follow a normal (Gaussian) distribution, with standard deviations of 1 cm for the LiDAR and 3 cm for the ultrasound sensor. Our objective is to combine (fuse) these measurements to produce a more accurate estimate.

A) Write down the formula for the CPDs $P(L \mid D)$ and $P(U \mid D)$.
B) Derive $P(D \mid L, U)$ and demonstrate that it also follows Gaussian distribution.
C) Give a better estimation as the maximizer of $P(D \mid L, U)$. Compare the error of this new estimation with those of LiDAR and ultrasound in terms of standard deviation. In other words, compare the standard deviation of $P(D \mid L, U)$ with those of $P(L \mid D)$ and $P(U \mid D)$.

Note: To do this you need to know $P(D)$. Not having any information about $P(D)$, you may assume that it has the uniform distribution $P(D) = \epsilon$ for $D \in [-\epsilon/2, \epsilon/2]$ and $P(D) = 0$ otherwise. Then consider the solution at the limit $\epsilon \rightarrow \infty$.