CSC412/2506 – Assignment #1

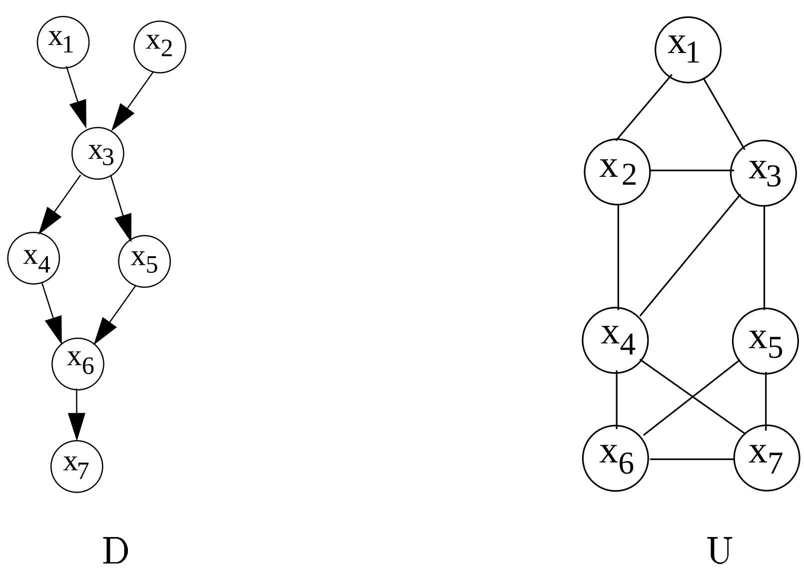
Due: Jan19, 10am at the **START** of class
Worth: 10%
Late assignments not accepted.

1 Graphical Model Distributions

The figure below shows a directed graphical model (D) and an undirected graphical model (U) each representing a joint distribution over six random variables.

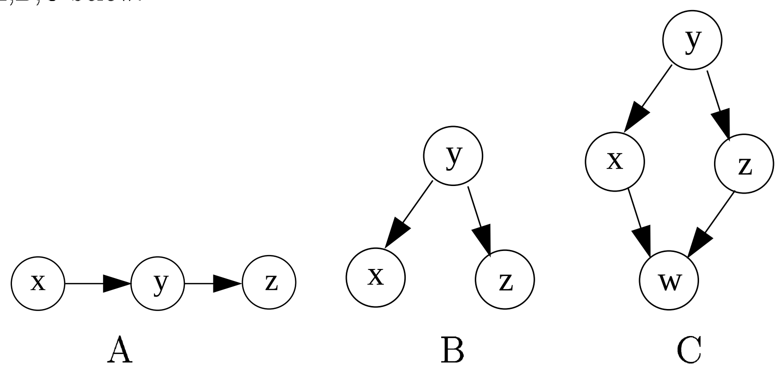
- For each of the following statements and each graphical model, prove whether the statement *must* be true of the distribution represented by the model, *could be true* but we don't know, or *cannot be true*.

  1. $\mathbf{x}_4$ is conditionally independent of $\mathbf{x}_5$ given $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$
  2. $\mathbf{x}_3$ is conditionally *dependent* on $\mathbf{x}_7$ given $\mathbf{x}_4$ and $\mathbf{x}_5$
  3. $\mathbf{x}_1$ is marginally independent of $\mathbf{x}_2$
  4. $\mathbf{x}_4$ is conditionally independent of $\mathbf{x}_5$ given $\mathbf{x}_3, \mathbf{x}_6$



2 Directed Models

For each of the directed graphs A,B,C below:



- Give all possible *topological orderings* of the variables.
- List all conditional independencies implied by the directed model.
- If there is an undirected model which captures exactly the same conditional independencies, draw it.
  If not, remove the fewest possible edges from the directed model until it has an undirected model capturing the equivalent independencies, and show the new directed model and the corresponding undirected model.

[Hard] Prove or disprove the following statement: If both the original directed graph and the undirected graph obtained by ignoring the direction of the edges are acyclic then there is an undirected model which captures exactly the same conditional independence assumptions as the directed model.

3 Complete Graphs

- Draw a directed graphical model on 5 variables which (a) can capture *any* joint distribution and (b) is acyclic.
- Can any edges be added to or removed from your graph and still preserve both the properties (a) and (b) above? If so, show the addition or removal, if not say why not.
- Draw an undirected graphical model on 4 variables which can capture *any* joint distribution.
  List all the maximal cliques.
- Can any edges be added to or removed from your graph and still preserve the above property?
  If so, show the addition or removal. If not, say why not.
- [Bonus] Prove or disprove the following statement: all directed acyclic graphical models on k variables which can capture any joint distribution have exactly $k^2/2 - k/2$ edges.
- [Extra Bonus] Prove or disprove the following statement: all undirected graphical models on k variables which can capture any joint distribution have exactly $k^2/2 - k/2$ edges.

4 Incompatible Conditionals

In this question you'll try to convince yourself that the representation

$$p(\mathbf{X}) = \prod_i \mathsf{P}(\mathbf{x}_i|\mathbf{x}_{\text{neighbours}(i)})$$

does not allow arbitrary conditional probabilities $\mathsf{P}(\mathbf{x}_i|\mathbf{x}_{\text{neighbours}(i)})$.

- Consider two binary variables $\mathbf{x}$ and $\mathbf{y}$ with the conditional distributions:

$$p(x=1|y=0) = 3/4 \qquad p(x=1|y=1) = 2/3$$
$$p(y=1|x=0) = 2/3 \qquad p(y=1|x=1) = 4/7$$

- Show that the function $f(x, y) = p(x|y)p(y|x)$ is not a valid joint distribution $p(x, y)$.
- Write down a valid joint distribution $p(x, y)$ that has the given conditionals.
- What condition must the joint $p(x, y)$ satisfy in order to be able to be written as a product of the conditionals p(x—y) and p(y—x)?
Show your reasoning.
- Write down two conditional distributions, like the ones above, on binary variables x and y, so that $p(x|y)p(y|x)$ *is* a valid joint. (Don't set any probabilities to zero or one.)

5 Numeric Distributions

This question is to get you comfortable with thinking about distributions as multidimensional tables, a concept which will be very useful for future assignments that include programming.

- Consider three multidimensional arrays (tables) `pA,pB,pC`, each representing a joint probability distribution over discrete random variables. Each dimension of the table represents one discrete random variable, and the size of the dimension is equal to the number of possible values the variable can take on. For example `pA(3,4,5,1)` is the probability under distribution A that $\mathbf{x}_1$ takes on its third value, $\mathbf{x}_2$ takes on its fourth value, $\mathbf{x}_3$ takes on its fifth value and $\mathbf{x}_4$ takes on its first value.
- Using the operations "select slice i along dimension j", "sum along the kth dimension" and "divide each element in the table by a scalar", describe, in order, the numerical operations that would be necessary to calculate the following marginal and conditional distributions and say what the size of the output array (table) would be if the original tables have size 5x5x5x5.
  1. $pA(\mathbf{x}_3, \mathbf{x}_4|\mathbf{x}_1$ takes its third value)
  2. $pB(\mathbf{x}_4|\mathbf{x}_2$ takes its first value)
  3. $pC(\mathbf{x}_1)$
- How would you decide if each of the following statements was true, given a multidimensional array representing a joint distribution?
  1. $\mathbf{x}_1$ is conditionally independent of $\mathbf{x}_2$ given $\mathbf{x}_3$
  2. $\mathbf{x}_1$ is conditionally independent of $\mathbf{x}_2$ given $\mathbf{x}_3$ and $\mathbf{x}_4$
  3. $\mathbf{x}_1$ is marginally independent of $\mathbf{x}_2$