CS 188
Fall 2020

Regular Discussion 7

Probability

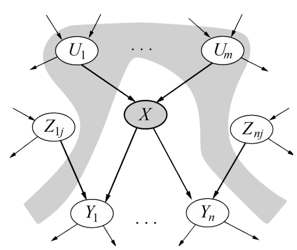
A **random variable** represents an event whose outcome is unknown. A **probability distribution** is an assignment of weights to outcomes A **joint distribution** over discrete random variables is a table of probabilities which captures the likelihood of each possible **outcome**, also known as an **assignment** of values to the random variables.

To write that random variables X and Y are **marginally independent**, we write $X \perp\!\!\!\perp Y$. To write that random variables X and Y are **conditionally independent** given another random variable Z, we write $X \perp\!\!\!\perp Y|Z$.

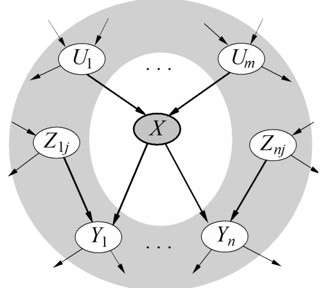
Bayesian Network Representation

In a Bayesian network, rather than storing information in a giant table, probabilities are instead distributed across a large number of smaller local probability tables along with a **directed acyclic graph** (DAG) which captures the relationships between variables. Thus, if we have a node representing variable X, we store $P(X|A_1, A_2, ..., A_N)$, where $A_1, ..., A_N$ are the parents of X.

- **Each node is conditionally independent of all its ancestor nodes (non-descendents) in the graph, given all of its parents**.

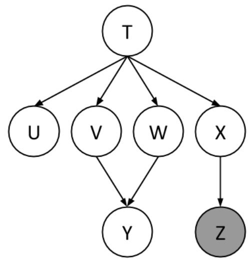


- **Each node is conditionally independent of all other variables given its Markov blanket.** A variable's Markov blanket consists of parents, children, children's other parents.



1 Variable Elimination

Using the same Bayes Net (shown below), we want to compute $P(Y \mid +z)$. All variables have binary domains. Assume we run variable elimination to compute the answer to this query, with the following variable elimination ordering: X, T, U, V, W.



Complete the following description of the factors generated in this process:

After inserting evidence, we have the following factors to start out with:

$$P(T), P(U|T), P(V|T), P(W|T), P(X|T), P(Y|V,W), P(+z|X)$$

(a) When eliminating X we generate a new factor f_1 as follows, which leaves us with the factors:

$$f_1(+z|T) = \sum_x P(x|T)P(+z|x) \qquad P(T), P(U|T), P(V|T), P(W|T), P(Y|V,W), f_1(+z|T)$$

(b) When eliminating T we generate a new factor f_2 as follows, which leaves us with the factors:

(c) When eliminating U we generate a new factor f_3 as follows, which leaves us with the factors:

(d) When eliminating V we generate a new factor f_4 as follows, which leaves us with the factors:

(e) When eliminating W we generate a new factor f_5 as follows, which leaves us with the factors:

(f) How would you obtain $P(Y \mid +z)$ from the factors left above:

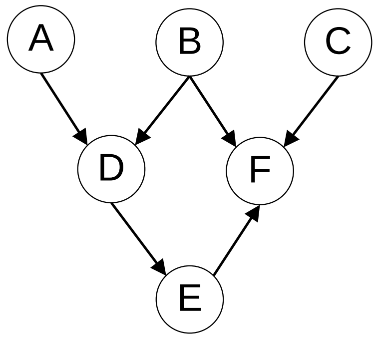
(g) What is the size of the largest factor that gets generated during the above process?

(m) Does there exist a better elimination ordering (one which generates smaller largest factors)?

2 Bayes Nets

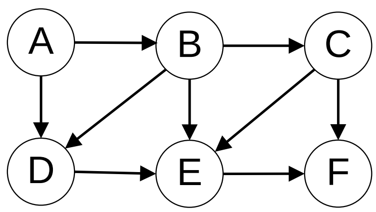
(a) For the following graphs, explicitly state the minimum size set of edges that must be removed such that the corresponding independence relations are guaranteed to be true.

Marked the removed edges with an 'X' on the graphs.



$$A \perp\!\!\!\perp B | F$$
$$A \perp\!\!\!\perp F | D$$
$$B \perp\!\!\!\perp C$$

(i)



$$A \perp\!\!\!\perp D | B$$
$$A \perp\!\!\!\perp F | C$$
$$C \perp\!\!\!\perp D | B$$

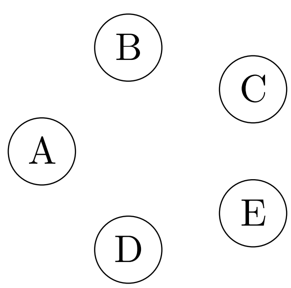
(ii)

(b) You're performing variable elimination over a Bayes Net with variables A, B, C, D, E. So far, you've finished joining over (but not summing out) C, when you realize you've lost the original Bayes Net!

Your current factors are $f(A), f(B), f(B, D), f(A, B, C, D, E)$. Note: these are factors, NOT joint distributions. You don't know which variables are conditioned or unconditioned.

(i) What's the smallest number of edges that could have been in the original Bayes Net? Draw out one such Bayes Net below.

Number of edges =



(ii) What's the largest number of edges that could have been in the original Bayes Net? Draw out one such Bayes Net below.

Number of edges =