Probabilistic Reasoning

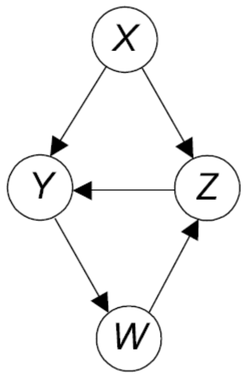
Unit # 6

Directed Graph

- A **directed graph** is a pair (V, E), where V is a finite, nonempty set whose elements are called **nodes** (or vertices), and E is a set of ordered pairs of distinct elements of V.
- Elements of E are called **directed edges**, and if $(X, Y) \in E$, we say there is an edge from X to Y.

Directed Graph (Cont'd)

- V = {X, Y, Z, W}
- E = {(X, Y), (X, Z), (Z, Y), (Y, W), (W, Z)}



Path and Cycle

- A **path** in a directed graph is a sequence of nodes $[X_1, X_2, ..., X_k]$ such that $(X_{i-1}, X_i) \in E$ for $2 \leq i \leq k$.
- A **cycle** in a directed graph is a path from a node to itself.

Directed Acyclic Graph

- A directed graph G is called a **directed acyclic graph (DAG)** if it contains no cycle.

More Definitions

- Given a DAG G = (V, E) and nodes X and Y in V,
  - Y is called a **parent** of X if there is an edge from Y to X,
  - Y is called a **descendent** of X and X is called an **ancestor** of Y if there is a path from X to Y, and
  - Y is called a **non-descendent** of X if Y is not a descendent of X and Y is not a parent of X.

Markov Condition and Bayesian Network

- Suppose we have a joint probability distribution P of the random variables in some set V and a DAG G = (V, E).
- We say that (G, P) satisfies the **Markov condition** if for each variable X $\in$ V, X is conditionally independent of the set of all of its non-descendents (ND) given the set of all its parents (PA).
- If (G, P) satisfies the Markov condition, we call (G, P) a **Bayesian network**.

Theorem

- (G, P) satisfies the Markov condition (and thus is a Bayesian network) if and only if P is equal to the product of its conditional distributions of all nodes given their parents in G, whenever these conditional distributions exist.

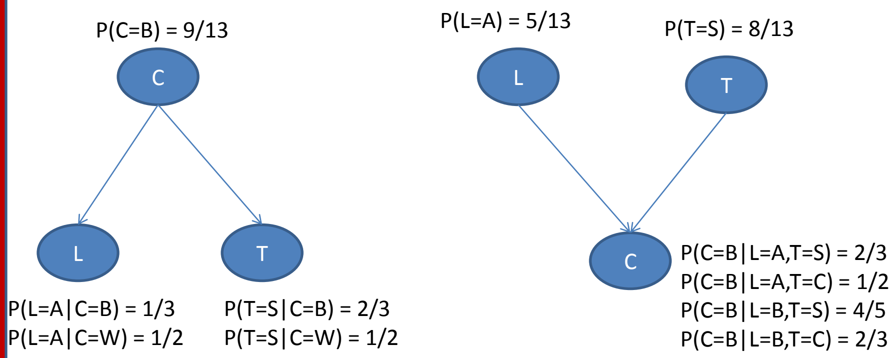
Caveat

- It is important to realize that we can't take just any DAG and expect a joint distribution to equal the product of its conditional distributions in the DAG.

Example Data

L	T	C
A	S	B
A	S	B
B	S	B
B	S	B
B	S	B
B	S	B
A	C	B
B	C	B
B	C	B
A	S	W
B	S	W
A	C	W
B	C	W

Example



- Which one of the above DAG satisfies the Markov condition?

Paradox

- Our goal is to represent a joint probability distribution succinctly using a DAG and conditional distributions for the DAG (a Bayesian network) rather than enumerating every value in the joint distribution.
- However, we don’t know which DAG to use until we check whether the Markov condition is satisfied, and in general, we would need to have the joint distribution to check this.

Paradox (Cont'd)

- A common way out of this predicament is to construct a causal DAG, which is a DAG in which there is an edge from X to Y if X causes Y.
- A second way of obtaining the DAG is to learn it from data.