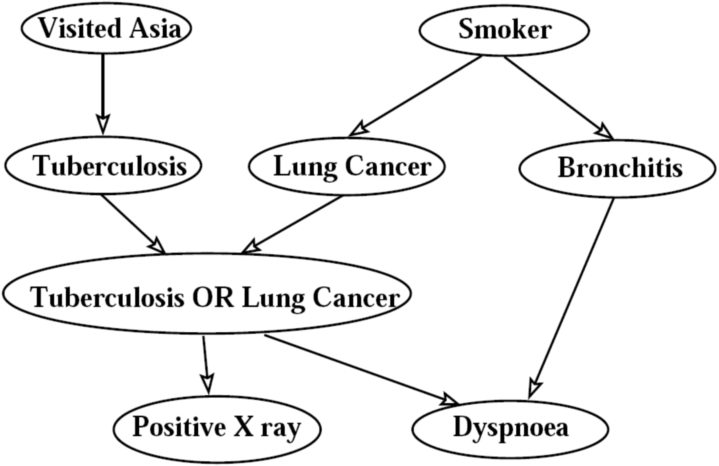
Graph Theory: Week 10

Probabilistic graphical models

John Quinn

23rd October, 2007

Probabilistic graphical models



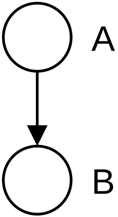
(Fig from David Barber, An Introduction to Graphical Models talk)

Probabilistic graphical models

- Directed acyclic graphs representing probability distributions.
- In directed case, also known as belief networks or Bayesian networks.
- Vertices represent random variables, arcs represent dependencies.
- The model is used to represent the *joint probability distribution* of all the variables.

Dependence

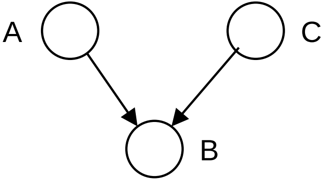
An arc from A to B implies that A “causes” B.



- A has no dependencies
- B is dependent on A.
- The two terms in this distribution are therefore P(A) and P(B|A).
- Joint distribution $P(A, B) = P(A)P(B|A)$

Conditional dependence

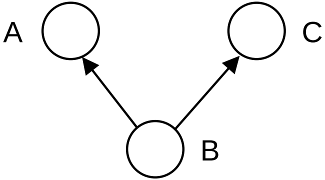
A and C are *conditionally dependent* given B.



- Imagine that B is "become rich", A is "inherit business" and C is "win competition".
- What if you knew someone was rich **and** that they had inherited a business? In the light of this information, would you think it was more likely or less likely that they had won the competition?
- Knowing A and B affects your belief in C, hence the dependency.

Conditional **in**dependence

A and C are *conditionally independent* given B.

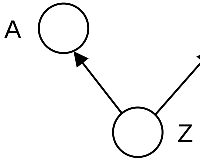


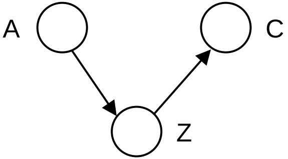
- Think of B as "being rich", A is "drive Mercedes" and C is "shop in Garden City".
- If you don't know B then A and C are dependent.
- If you do know B, then A provides no information about C and vice versa.

d-separation

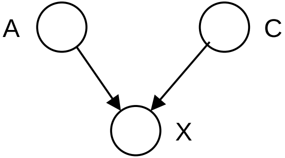
Two vertices A,C are d-separated by Z iff

- The following cases are true:





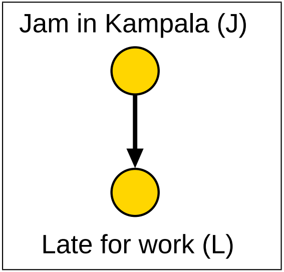
- Or the following:



  where neither X nor any of its descendants are in Z.
- If two vertices A and C are d-separated by a vertex Z, then A and B are conditionally independent on Z. That is, $P(A|B,Z) = P(A|Z)$ and vice versa.

Parameters of the model

Having the (in)dependency structure of the variables, need to specify what the particular distribution is for each vertex. In this case, the parameters are contained in conditional probability tables:



Probability of a jam

P(J=True)	0.4
P(J=False)	0.6

Probability of being late

P(L=True|J=True)	0.9
P(L=False|J=True)	0.1
P(L=True|J=False)	0.2
P(L=False|J=False)	0.8

Observed and latent variables

- Vertices in a graphical model can be divided into two categories: observed (measured) and latent (hidden).
- Some variables can be measured—we can find out what their values are.
- There are other variables for which we want to try and work out the value based on the things that we can measure.

Inference

- The most common use for graphical models is to make inferences.
- Can try to manipulate the joint probability using *marginalisation* and *Bayes rule* in order to obtain answers.
- In general the problem can be difficult and there are techniques to automate and approximate the process.

Marginalisation

Sum over parent vertices to eliminate them.

$$P(A) = \sum_b P(A|B = b)$$

Can use this e.g. in the traffic jam example to eliminate J when inferring L.

Models which represent time

- Dynamic applications—speech recognition, tracking, control. . .
- We can design graphical models which have vertices arranged to represent a sequence of time frames.