Title: Probabilistic Reasoning Systems (II)

AIMA: Chapter 15 (Section $\sim$ 15.3-15.4)

Introduction to Artificial Intelligence

CSCE 476-876, Spring 2000

www.cse.unl.edu/~cse476

Berthe Y. Choueiry

choueiry@cse.unl.edu, (402)472-5444

Inference in BN

- $\surd$ Exact inference:
  Propagate "messages" along the structure of the network
  — Enumeration
  — Variable elimination

- $\times$ Approximate inference: Stochastic techniques
  Fast but not exact
  — Stochastic simulation
  — Markov chain

Basic inference

Adapted from the lecture notes of Daphne Koller

Given: A and B two binary/bi-valued random variable such that..

$$A \longrightarrow B$$

Question: Compute $P(B)$

$$P(B) = P(A)P(B|A) + P(\neg A)P(B|\neg A)$$

From the CTP of A, we have: $P(A), P(\neg A)$
From the CTP of B, we have: $P(B|A), P(B|\neg A)$

If A and B were multi-valued variables:

$$\forall B = b,\ P(B = b) = \sum_i P(a_i)P(B = b|a_i)$$

with $P(a_i)$ available from the CPT of A,
and, $P(B = b|a_i)$ available from the CPT of B

Consider a longer chain:

$$A \longrightarrow B \longrightarrow C$$

Question: Compute $P(B)$

Same as before:

$$\forall C = c,\ P(C = c) = \sum_b P(b)P(C = c|b)$$

with $P(b)$ just computed before,
and, $P(B = b|a_i)$ available from the CPT of B

If C depended on A too, this would not have worked

Again, if we have:

$$A \longrightarrow B \longrightarrow C \longrightarrow D$$

.. we can do inference in linear time in #nodes in the chain

If each node has k values$\longrightarrow O(nk^2)$

Enumeration (JDT) $\longrightarrow O(k^n)$

So:

$$P(d) = \sum_{a,b,c} P(a,b,c,d)$$

By the chain rule:

$$P(d) = \sum_{a,b,c} P(a)P(b|a)P(c|b)P(d|c)$$

Opening up the summation:

$$P(d) = \sum_{c}\sum_{b}\sum_{a} P(a)P(b|a)P(c|b)P(d|c)$$

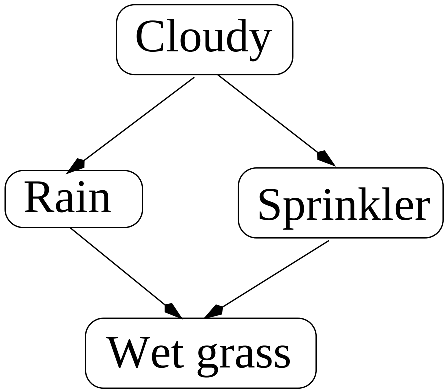
Pushing forward summations:

$$P(d) = \sum_{c} P(d|c) \sum_{b} P(c|b) \sum_{a} P(a)P(b|a)$$

Doing the procedure in the standard way, we would have to compute $P(b) = \sum_a P(a)P(b|a)$ many times, once for every value of C and D

In a chain: internal summations computed exponentially many times!

More complex example



$$
\begin{aligned}
P(w) &= \sum_{r,s,c} P(w|r,s)P(r|c)P(s|c)P(c) \\
&= \sum_{r,s} P(w|r,s) \sum_{c} P(r|c)P(s|c)P(c)
\end{aligned}
$$

First compute: $\sum_c P(r|c)P(s|c)P(c)$..

for all $c \longrightarrow$ computing entire CPTs

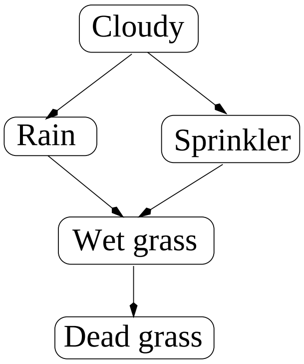
Variable elimination

> *Look at the expression of summing over the joint*
> *push each summation in as far as it will go*

The performance of the algorithm depends heavily on the elimination order (related to Gaussian elimination.. optimal orders that minimize the number of fill-in terms.. triangulated graphs.. *etc.*)

Question: How to handle evidence?

Question: Compute $P(c|d)$



$P(c|d) = \frac{P(c,d)}{P(d)}$

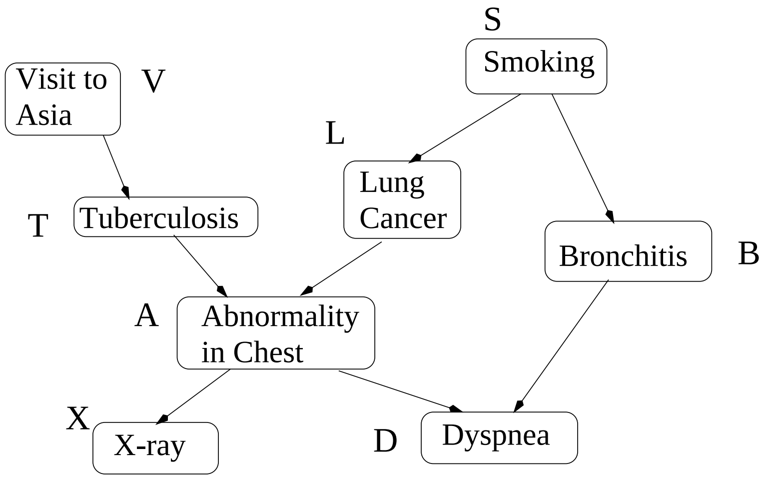
1- $P(c, d)$
$P(C, d) =$ sum over all entries in the JPD consistent with $C = c, D = d$

$$\begin{aligned} P(C,d) &= \sum_{w,r,s} P(c,r,s,w,d) \\ &= \sum_{w,r,s} P(c)P(s|c)P(r|c)P(w|r,c)P(d|w) \\ &= \sum_{w} P(d|w) \sum_{r,s} P(c)P(s|)P(r|c)P(w|r,c)P(d|w) \end{aligned}$$

Compute expression inside out..

2- $P(d) = \sum_c P(c, d)$

Question: Compute $P(d)$



$$P(d) = \sum_{a,b,l,t,s,x,v} P(v)P(t|v)P(s)P(l|s)P(b|s)P(a|l,t)P(x|a)P(d|a,b)$$

(1) Sum over V, **(2)** Sum over X, **(3)** Sum over S, **(4)** Sum over L, **(5)** Sum over B, **(6)** Sum over T, **(7)** Sum over A & B

Problem: NP-hard in the general case

The sparser, the easier

Special graphs (polytrees), easy

Inference in BN

Compute the probability of a variable
given the probabilities of some other variables

Three items are necessary:

- **Causal support**:
  Total support arriving at the node from its parent nodes (its causes)
- **Evidential support**:
  Total support arriving at the node from its children (its symptoms)
- The entry in the CPT relating the node to its causes

AIMA Exercise 15.1