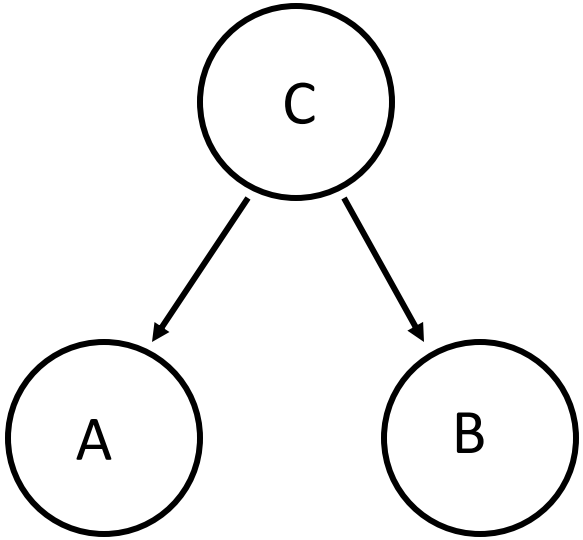
Dentist Example



A is conditionally independent of B given C

B is conditionally independent of A given C

Given C (cavity or no cavity)

A and B may or may not be true

but each is independent of the other

If we do not know C,

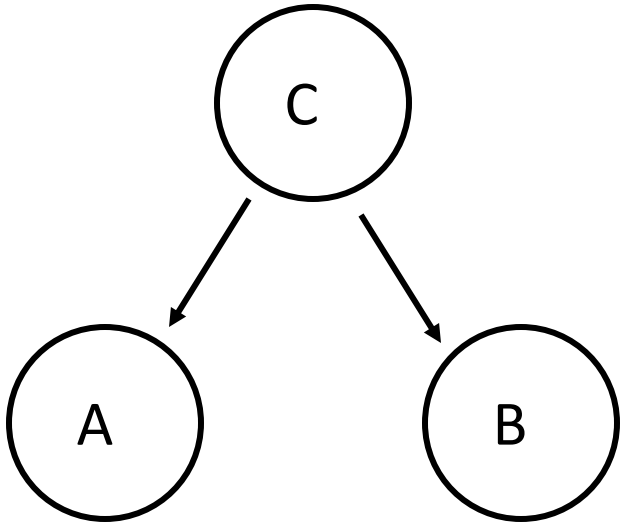
A influences our belief in B & vice versa

A is evidence for C which in turn changes our opinion of B

CPT
(conditional probability table)

- For each random variable (node)
- Table of distributions
- Each is a distribution over the possible values of the random variable
- For each *configuration* of its parents' values, there is one distribution for the R.V.s values
- Represents the probability of each R.V. value conditioned on its parents

How Many Parameters



Joint:

$2^3 - 1 = 7$ (why?)

Bayes Net:

Distribution(s) at C

Pr(cavity), Pr(¬cavity)

Need one parameter

Distribution(s) at A

Pr(ache|cavity), Pr(¬ache|cavity)
Pr(ache| ¬cavity), Pr(¬ache| ¬cavity)

Distributions at B, like A

1 + 2 + 2 = 5 (a savings of 2!)

A BN represents the full Joint

(it contains all of the information in the Joint)

- The elements of the joint are recoverable from the BN

$$P(x_1, x_2, x_3, \ldots x_n) = \prod_{i=1}^{n} P(x_i \mid parents(X_i)) \quad \text{(eqn 14.2)}$$

- The Joint is "factorable"
- Can we represent all possible Joints for a BN such as the Dentist example?

How flexible is
The Joint
vs.
A Bayesian Net

Q: What Joints cannot be represented?

A: Those that violate the BNs conditional independence assumptions

Q: How is a conditional independence assumption denoted in a BN?

A: The *lack* of an arc.

Where do the CPT entries come from?

- Make them up
  (since we are Bayesians)
- Count
  (count what?)

Dentist Example

Boolean RV's C: cavity, A: ache, B: probe catches

CPT's

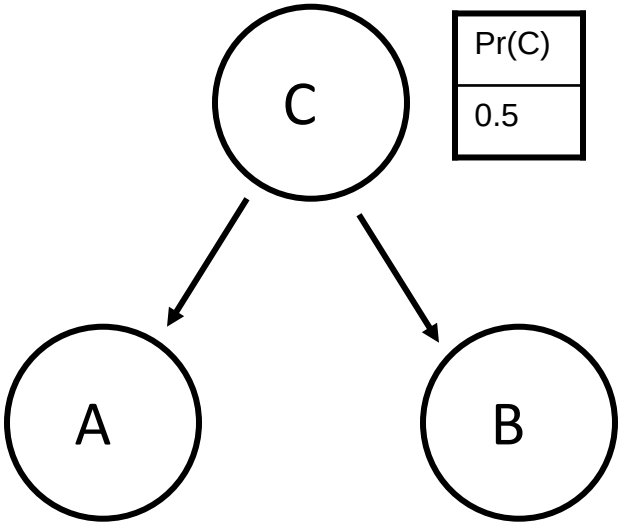
At C: Pr(C)

At A: Pr(A | C)

At B: Pr(B | C)

Training Examples:

Person	Cavity	Ache	Probe
1	T	F	T
2	F	T	F
3	T	T	T
4	F	F	F
5	T	T	T
6	T	F	T
7	F	F	T
8	F	F	F



Pr(C)
0.5

C	Pr(A|C)
T	0.5
F	0.25

C	Pr(B|C)
T	1.0
F	0.25

These are maximum likelihood estimates

Really need more data

Probably should smooth

What do these numbers mean?

- For Dentist BN they summarize a Beta distribution
- For more outcomes…
  a Dirichlet distribution
- Why a *summary*?
- The standard CPT loses confidence information
- Pr(ache | cavity) = 0.5
  - because 2 of 4 cavities reported ache
  - what if it were ~ 30 of 60
  - or ~ 7,000 of 14,000?

Inference in Bayes Nets

- What is the probability of some configuration (the query)
- Given:
  - A distribution
  - Some observations (the evidence)
- In our BN
  - What is the probability that John has a cavity and no toothache (the query)
  - Given that the dentist probe catches (the evidence)

- Query variables; Evidence variables; Nuisance Variables