Course Logistics

- Any suggestions on streamlining homework / MPs?
- Next: BN MP (watch for it)

Inference

- Realizing what you already know
  (Inferencing does not make things true)
- Logical inference – propagating truth
  - Given some axioms (believe these)
  - Derive theorems (must believe these)
- Statistical inference – propagating evidence
  - Given some distributions, observations, etc.
  - Estimate other distributions, parameter values, probabilities, etc.
  - Uncertainty is fundamental

Conditional Independence

- Model the disease influencing the symptoms
- But no symptom interactions *given* the disease
- Conditional independence:

$$P(A|B,C) = P(A|C)$$

"A is conditionally independent of B given C"

Bayesian Networks
(belief networks,…)

- Graphical model
  - Probability + Graph theory
- Nodes – discrete random variables
- Directed Arcs – direct influence
  - Usually “causality” (class, text,…)
- Configuration of parents establishes contexts for a node
  - Different distribution for each context
  - Conditional Probability Tables (CPTs)
- No directed cycles (DAG)
- General graphical models
  - Markov Networks
  - Conditional random fields
  - Dynamic Bayesian nets
  - others

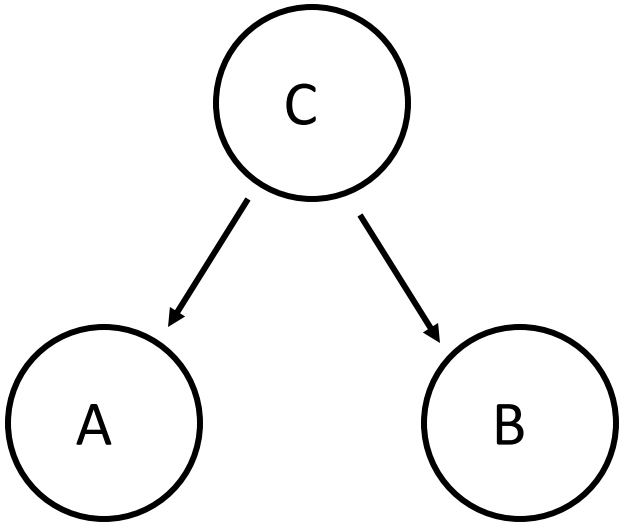
Dentist Example

3 Boolean Random Variables:

C – Patient has a cavity

A – Patient reports a toothache

B – Dentist's probe catches on tooth



C directly influences A and B

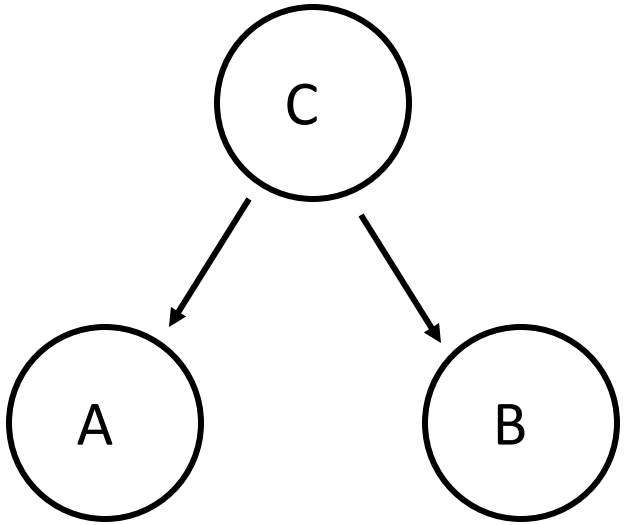
A and B are conditionally independent given C

CPT at each node specifies a probability distribution for each context (configuration of parents)

How many numbers do we need for the full joint?

How many for the Bayesian net?

How Many Numbers



Joint:

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

Bayes Net:

Distribution(s) at C

P(cavity), P(¬cavity)

Need one number

Distribution(s) at A

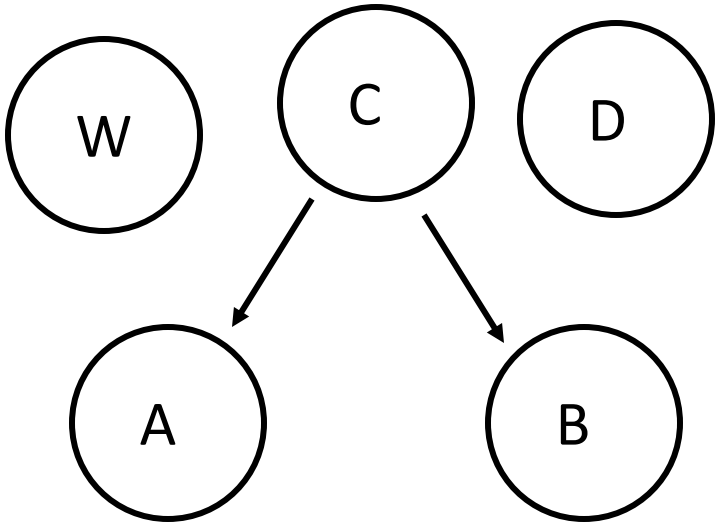
P(ache|cavity), P(¬ache|cavity)
P(ache| ¬cavity), P(¬ache| ¬cavity)

Distributions at B, like A

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

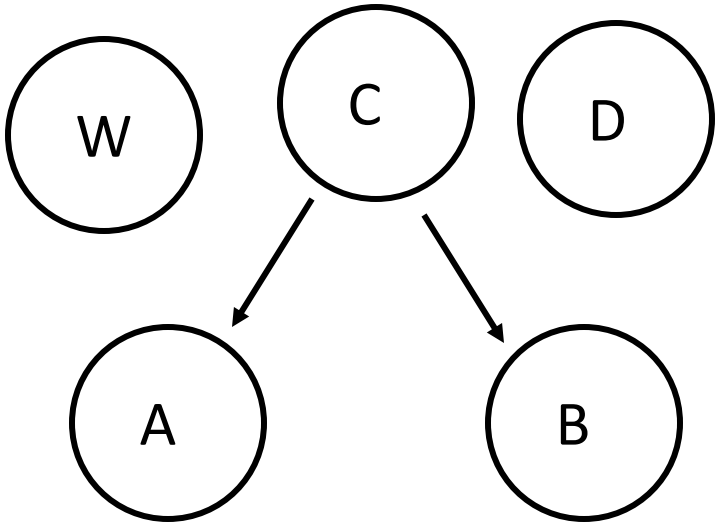
How Many Numbers?

- Add binary random variables for
  - Weather W
  - Weekday / Weekend D
- How many numbers for the joint? 31
- How many numbers for the BN? 7
- What does the BN look like?



How Many Numbers?

- Suppose A,B,C,D,W are trinary
- How many numbers for the joint? $3^5 - 1 = 242$
- How many numbers for the BN? 2+2+2+6+6 = 18
- Suppose A is binary, B is trinary, C takes on 4 values D takes on 5, W takes on 6?



W 5, D 4, C 3, A 4, B 8 = 24

Notation

- W, C, D, A, B are random variables
  - C (binary presence of cavity) ranges over values
  - C's values might be Yes or No; 1 or 0; c_1 or c_2

- Probability distribution: P(C)
- Probability value: $P(C=c_1)$ or $P(c_1)$ or P(c)

- P(A,B,C,W,D) full joint probability distribution
- P(a,b,c,w,d) some particular entry in the joint

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"
- Use any topological ordering of the nodes
- Can we represent all Joints over $X_1, X_2, \ldots X_n$?
- Which cannot be represented in the BN?

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?

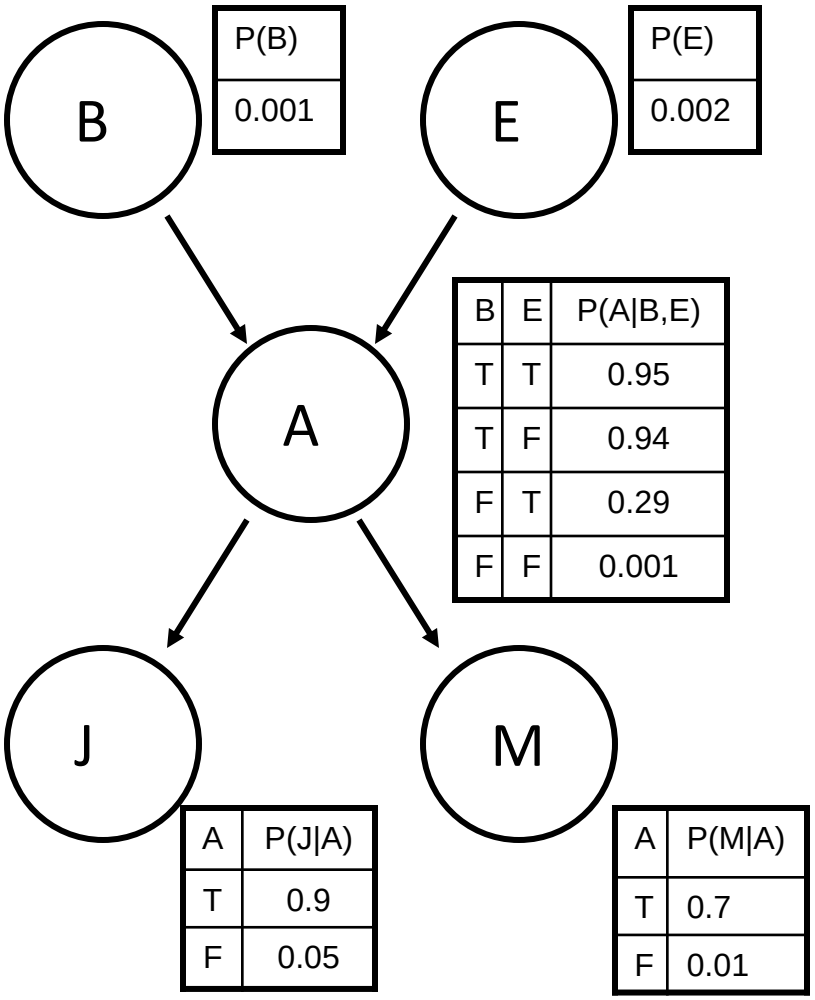
A: The *lack* of an arc.

A Statistical Model
(for us in AI)

- A (compact) representation of a particular statistical distribution
- Structural Part:
  - Commitment to random variables
    - Observable
    - Latent (Unobservable)
  - Captures restrictions on how they can interact
- Parametric Part:
  - Set of parameters (sufficient statistics)
  - Their values individuate a distribution

Bayes Net Model

Structural Part



Parametric Part

Where do the CPT numbers come from

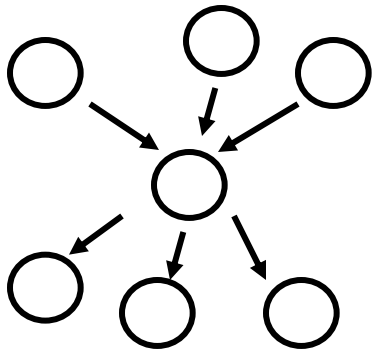
- Count
- Make them up
  (since we are Bayesians)

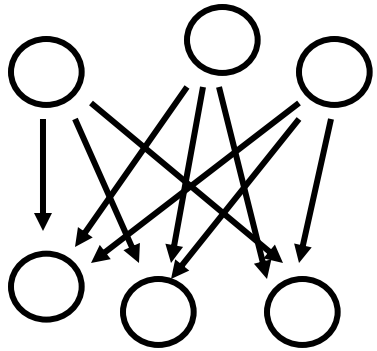
Where does the BN structure come from?

- Make it up
  (elicit it from experts)
- Learn from data
  - Works for special classes of nets
  - Can require too much data
  - Can be ill-posed

Missing Values in Observations

- Noise – unsystematically erased
- Latent
  - Values never observed
  - Ex: “applicant honesty” in credit card form
  - Influences other variables





How many parameters (if Boolean)?

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

Inference with BNs

Five Boolean Random Variables:

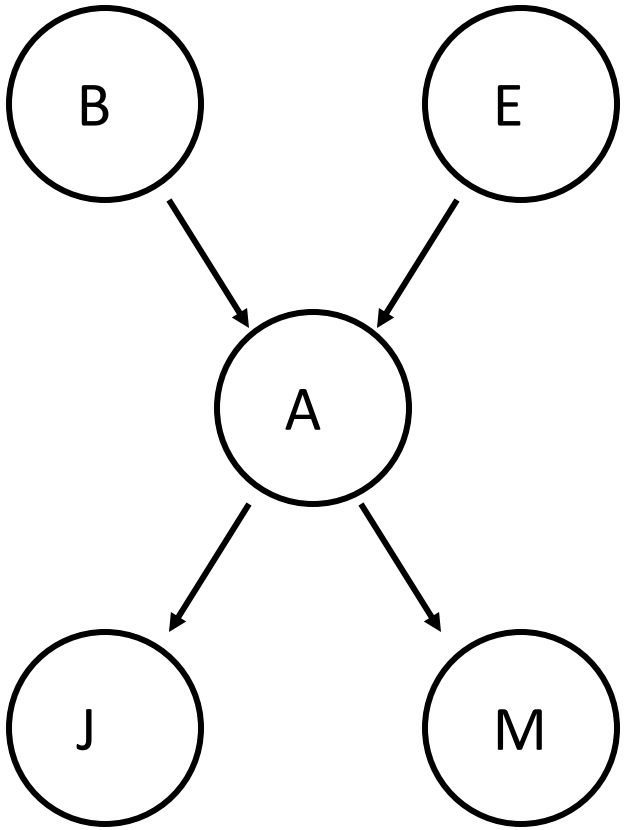
B – a burglary is in progress

E – an earthquake is in progress

A – the alarm is sounding

J – John calls

M – Mary calls



What's missing?

Inference with BNs

Five Boolean Random Variables:

B – a burglary is in progress

E – an earthquake is in progress

A – the alarm is sounding

J – John calls

M – Mary calls

Are these numbers reasonable?

How would they be different if they were Joint entries?

We can compute Joint entries.