Reasoning Under Uncertainty: Belief Networks

CPSC 322 Lecture 26

R&R systems we'll cover in this course

		Environment	
Problem		Deterministic	Stochastic
Static	Constraint Satisfaction	*Variables + Constraints* Search Arc Consistency Local Search	
	Query	*Logics* Search	*Bayesian (Belief) Networks* Variable Elimination
Sequential	Planning	*STRIPS* Search	*Decision Networks* Variable Elimination

Representation
Reasoning Technique

Key points Recap

We model the **environment** as a set of ….

Why is the **joint distribution** not an adequate representation ?

- "Representation, reasoning and learning" are **exponential** in …..

Solution: Exploit marginal & **conditional** independence

$$P(X|Y) = P(X) \qquad P(X|YZ) = P(X|Z)$$

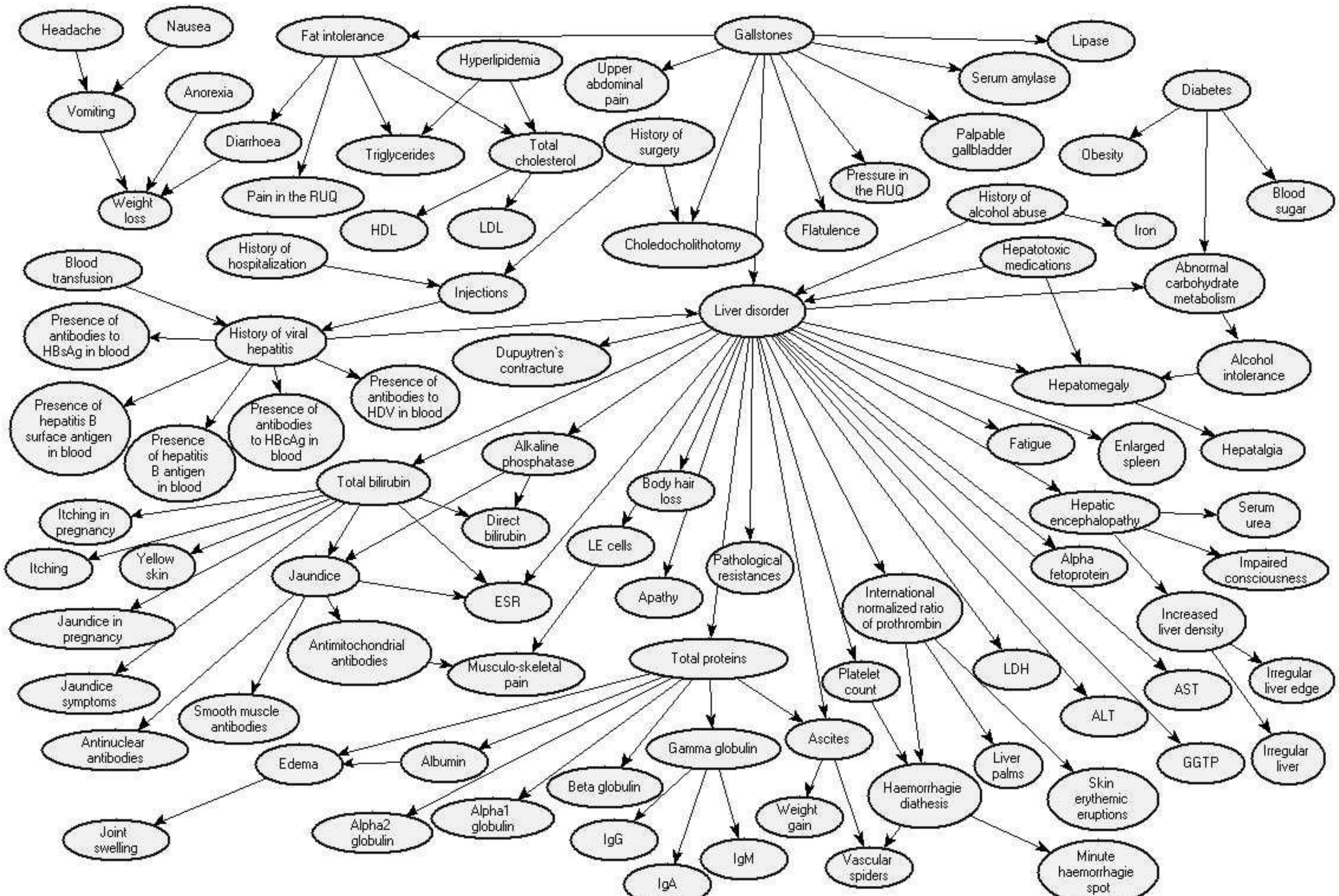
But how does independence allow us to simplify the joint?

CHAIN RULE!

Realistic BNet: Liver Diagnosis

Source: Onisko et al., 1999

~60 binary variables: how large is the JPD?



Learning Goals for today's class

You can:

Build a Belief Network for a simple domain

Classify the types of inference

Compute the representational saving in terms on number of probabilities required

Lecture Overview

- **Belief Networks**
  - **Build sample BN**
  - Intro Inference, Compactness, Semantics
  - More Examples

Belief Nets: Burglary Example

There might be a **Burglar** in my house

The **anti-burglar Alarm** in my house may go off

I have an agreement with two of my neighbors, **John** and **Mary**, that they **call** me if they hear the alarm go off when I am at work

Minor Earthquakes may occur and sometimes the set off the alarm.

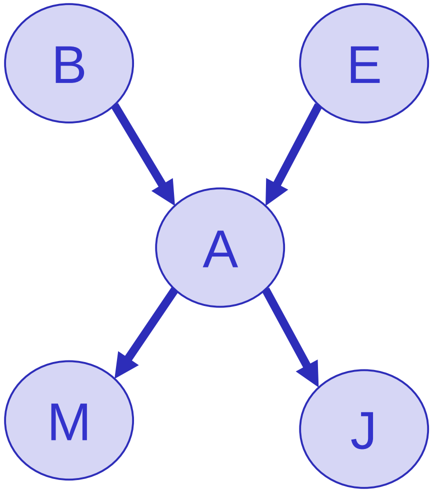
Variables: B, A, J, M, E

Joint has entries/probs

Belief Nets: Simplify the joint

- Typically order vars to reflect causal knowledge (i.e., causes *before effects*)
  - A burglar (B) can set the alarm (A) off
  - An earthquake (E) can set the alarm (A) off
  - The alarm can cause Mary to call (M)
  - The alarm can cause John to call (J)

$$P(B, E, A, M, J)$$



- Apply Chain Rule

$$P(B)\ P(E|B)\ P(A|B,E)\ P(M|A,E,B)P(J|M,A,E,B)$$

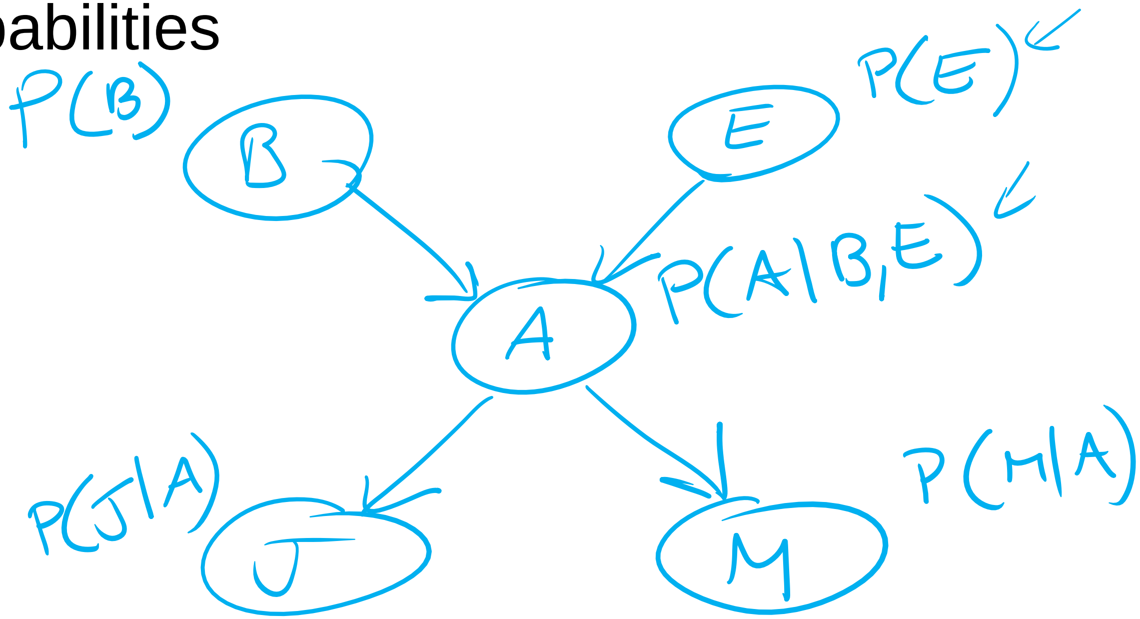
- Simplify according to marginal & conditional independence

$$P(B)\ P(E)\ P(A|B,E)\ P(M|A)\ P(J|A)$$

Belief Nets: Structure + Probs

$$P(B) * P(E) * P(A|B,E) * P(M|A) * P(J|A)$$

- Express remaining dependencies as a network
  - Each var is a node
  - For each var, the conditioning vars are its parents
  - Associate to each node corresponding conditional probabilities



- Directed Acyclic Graph (DAG)

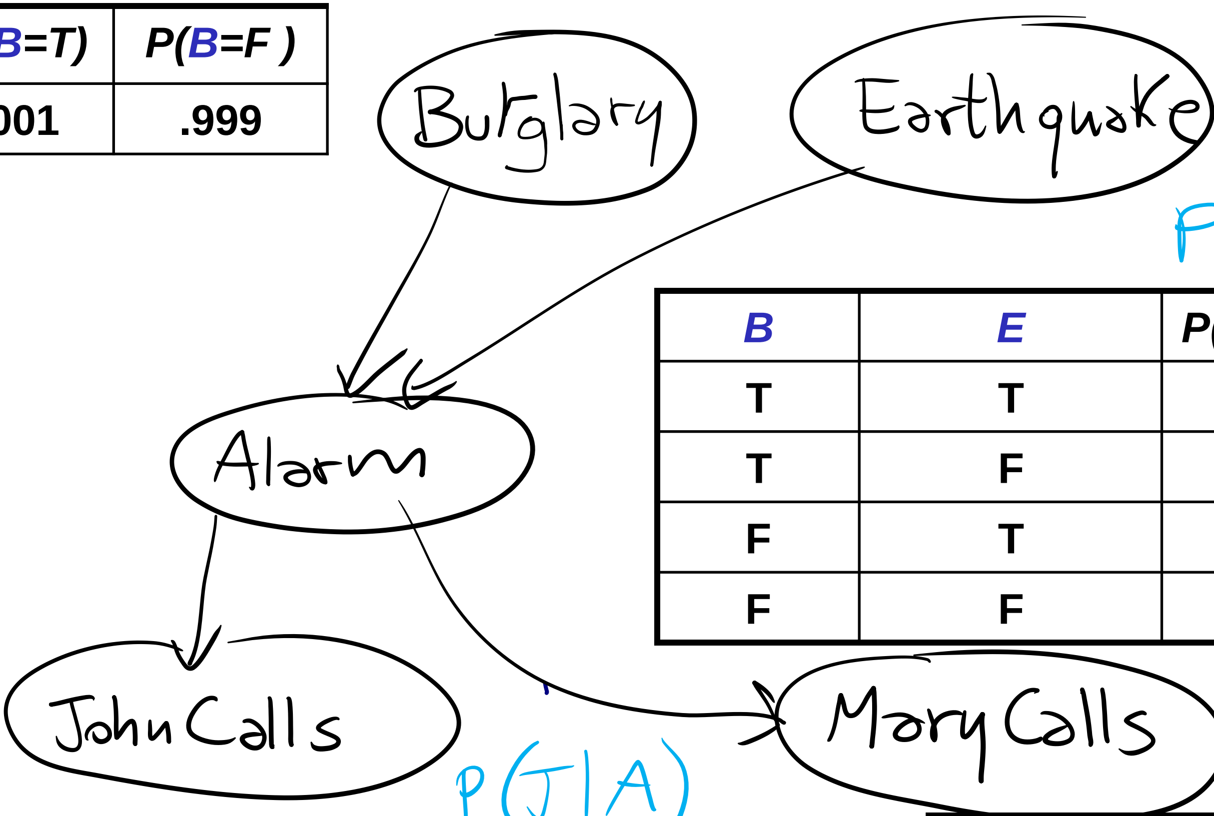
Burglary: complete BN

P(B)

P(B=T)	P(B=F)
.001	.999

P(E)

P(E=T)	P(E=F)
.002	.998



P(A | B,E)

B	E	P(A=T | B,E)	P(A=F | B,E)
T	T	.95	.05
T	F	.94	.06
F	T	.29	.71
F	F	.001	.999

P(J | A)

A	P(J=T | A)	P(J=F | A)
T	.90	.10
F	.05	.95

P(M | A)

A	P(M=T | A)	P(M=F | A)
T	.70	.30
F	.01	.99

call for any other reasons

Lecture Overview

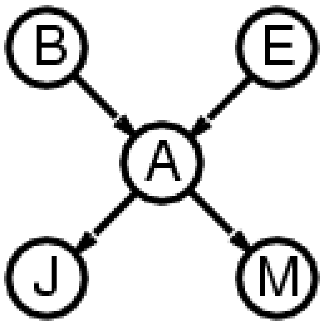
- **Belief Networks**
  - **Build sample BN**
  - **Intro Inference, Compactness, Semantics**
  - More Examples

Burglary Example: Bnets inference

Our BN can answer any probabilistic query that can be answered by processing the joint!

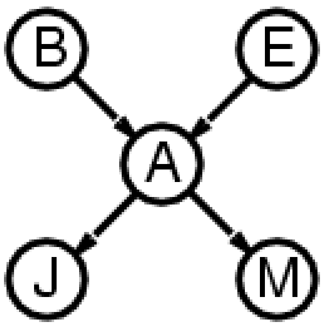
(Ex1) I'm at work,

- neighbor John calls,
- neighbor Mary doesn't call.
- No news of any earthquakes.
- Is there a burglar?



(Ex2) I'm at work,

- Receive message that neighbor John called ,
- News of minor earthquakes.
- Is there a burglar?



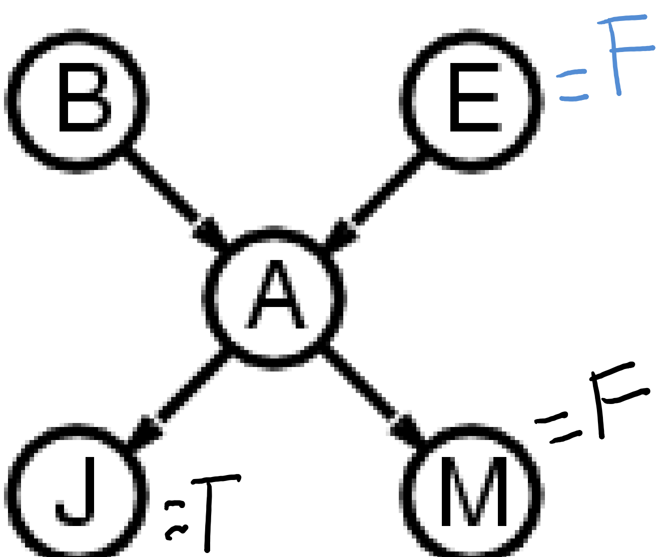
Set decimal places to monitor to 5

Burglary Example: Bnets inference

Our BN can answer any probabilistic query that can be answered by processing the joint!

(Ex1) I'm at work,

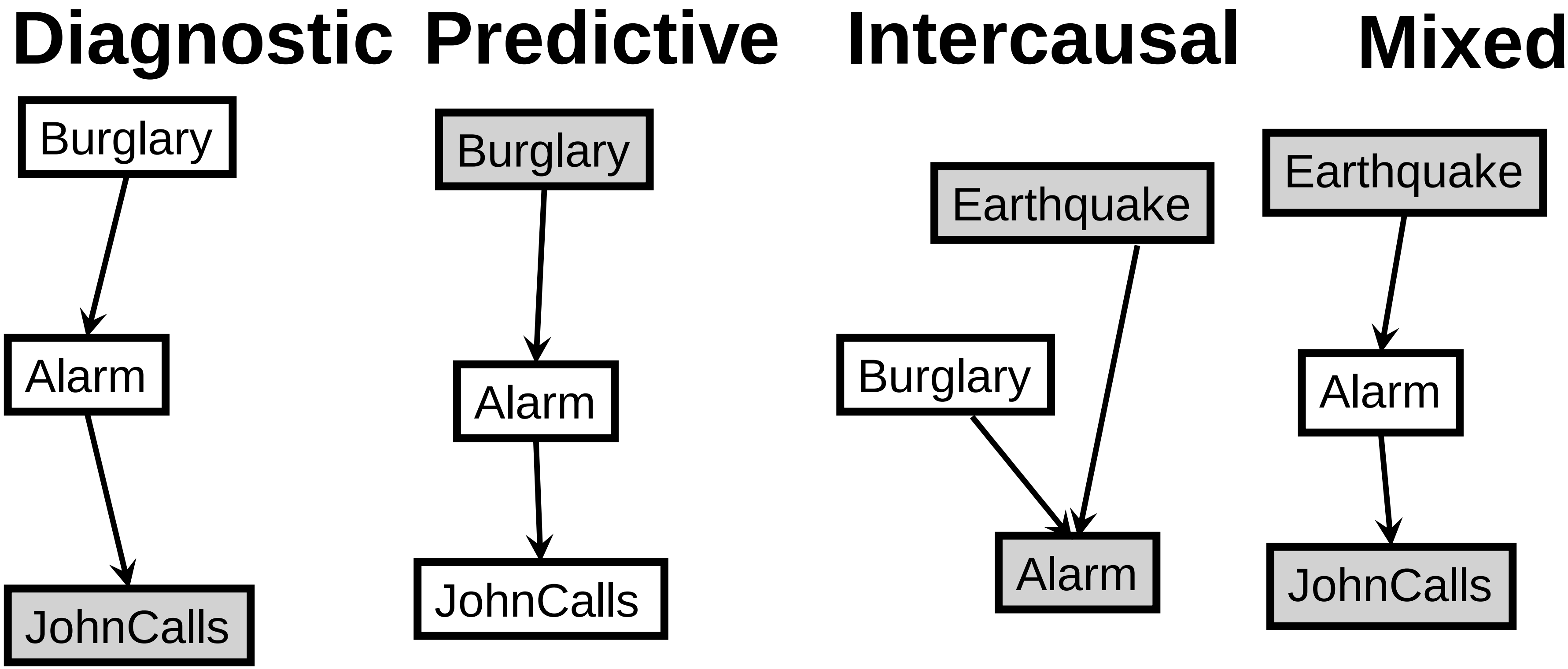
- neighbor John calls,
- neighbor Mary doesn't call.
- No news of any earthquakes.
- Is there a burglar?



The probability of Burglar will:

A. Go down
B. Remain the same
C. Go up

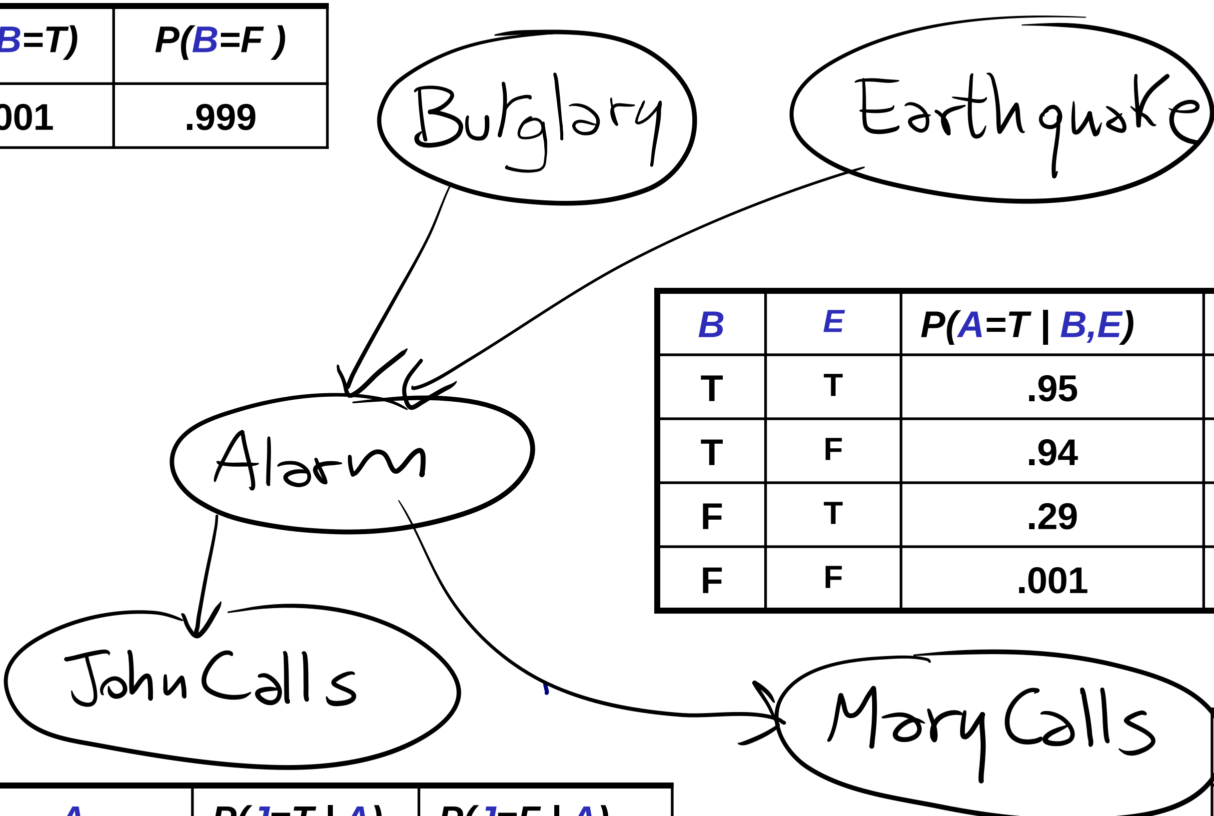
Bayesian Networks – Inference Types



BNnets: Compactness

P(B=T)	P(B=F)
.001	.999

P(E=T)	P(E=F)
.002	.998



B	E	P(A=T | B,E)	P(A=F | B,E)
T	T	.95	.05
T	F	.94	.06
F	T	.29	.71
F	F	.001	.999

A	P(J=T | A)	P(J=F | A)
T	.90	.10
F	.05	.95

A	P(M=T | A)	P(M=F | A)
T	.70	.30
F	.01	.99

$|JPD| = 2^5 - 1$

How many values do we **have to** store with this BNet?

A. 5 B. 10 C. 16 D. 20 E. 42

BNets: Compactness

A **Conditional Probability Table (CPT)** for **boolean** X_i with k **boolean** parents has 2^k rows for the combinations of parent values

Each row requires **one number** p_i for $X_i = true$
(the number for $X_i = false$ is just $1-p_i$)

If **each variable** has no more than ***k* parents**, the complete network requires $O(n2^k)$ numbers

For $k \ll n$, this is a substantial improvement,

- the numbers required grow linearly with n, vs. $O(2^n)$ for the full joint distribution

BNets: Construction General Semantics

The full joint distribution can be defined as the product of conditional distributions:

$$\boldsymbol{P}(X_1, \ldots, X_n) = \prod_i \boldsymbol{P}(X_i \mid X_1, \ldots, X_{i-1}) \text{ (chain rule)}$$

Simplify according to marginal/conditional independence

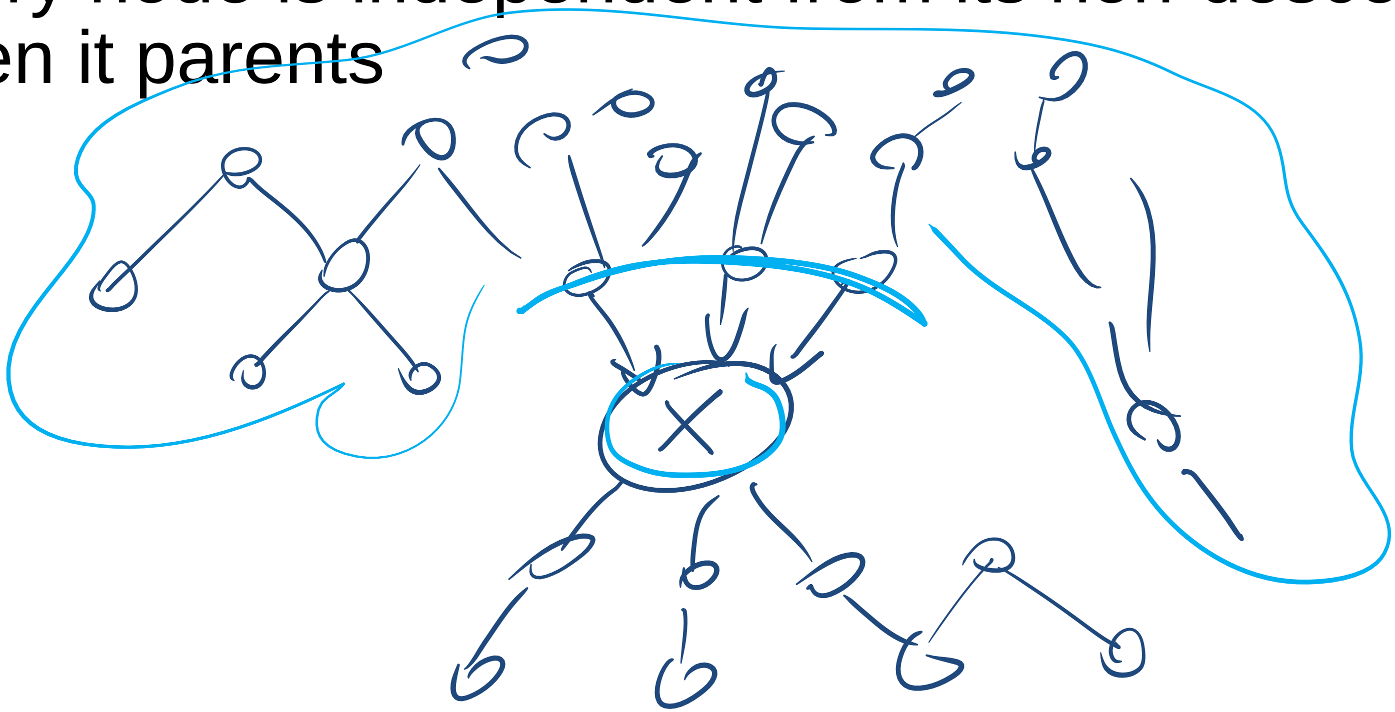
- Express remaining dependencies as a network
  - Each var is a node
  - For each var, the conditioning vars are its parents
  - Associate to each node corresponding conditional probabilities

$$\boldsymbol{P}(X_1, \ldots, X_n) = \prod_i \boldsymbol{P}(X_i \mid Parents(X_i))$$

BNets: Construction General Semantics (cont')

$$\mathbf{P}\,(X_1, \ldots, X_n) = \prod_{i=1}^{n} \mathbf{P}\,(X_i \mid Parents(X_i))$$

- Every node is independent from its non-descendants given it parents



(CPSC 422) A node is independent from the rest of the **network** given its **Markov Blanket**

Lecture Overview

- **Belief Networks**
  - **Build sample BN**
  - Intro Inference, Compactness, Semantics
  - **More Examples**

Other Examples: Fire Diagnosis (textbook 8.3.2)

Suppose you want to diagnose whether there is a fire in a building

- you receive a "noisy" report about whether everyone is leaving the building.
- if everyone is leaving, this may have been caused by a fire alarm.
- if there is a fire alarm, it may have been caused by a fire or by tampering
- if there is a fire, there may be smoke raising from the bldg.

In-class activity: build a Bayesian network for this environment and state what probability tables are needed

Other Examples (cont')

- Make sure you explore and understand the **Fire Diagnosis** example (we'll expand on it to study Decision Networks)

- **Electrical Circuit** example (textbook ex 6.11, 1st ed.)

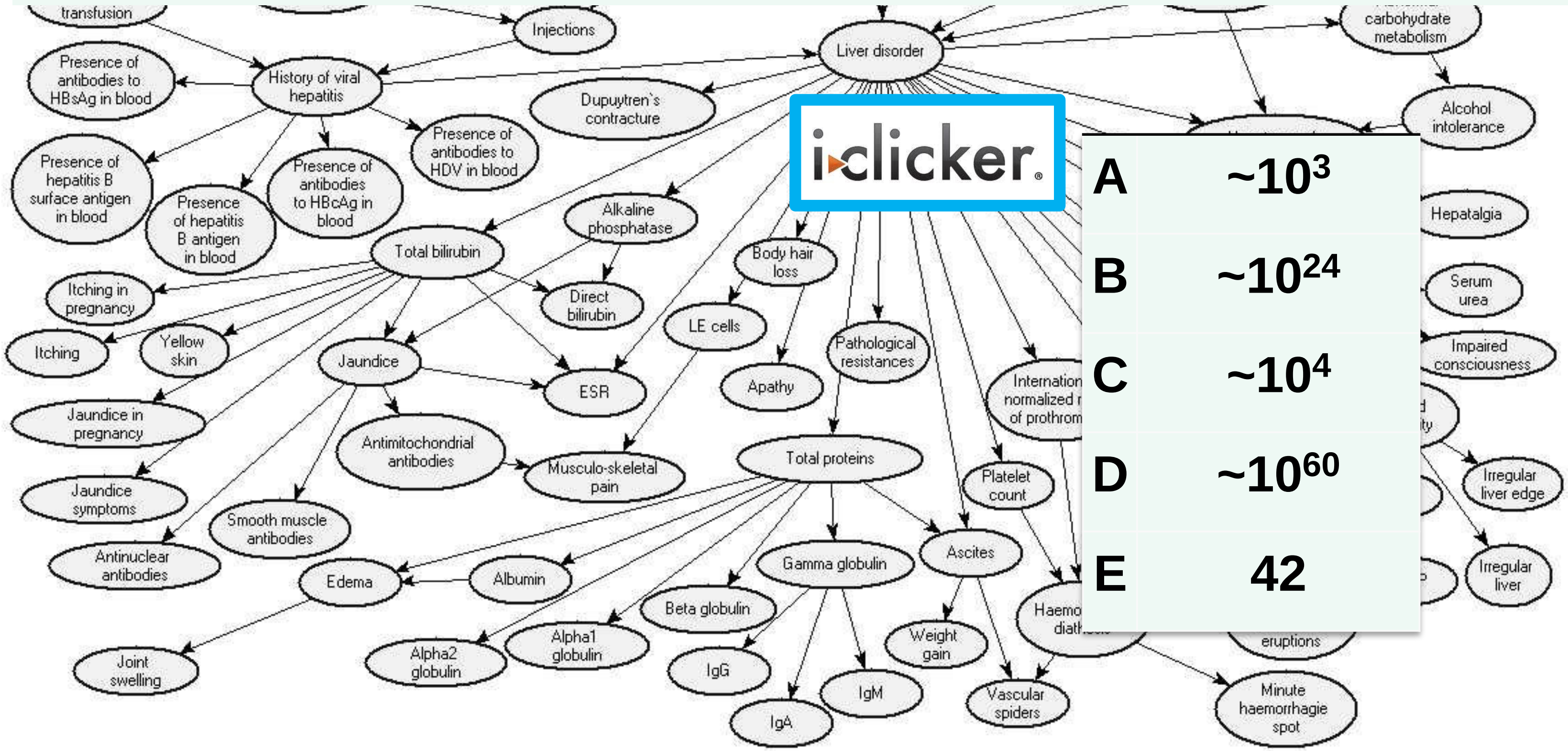
- **Patient's wheezing and coughing** example (ex. 6.14, 1st ed.)

- Other examples on

Realistic BNet: Liver Diagnosis

Source: Onisko et al., 1999

Assuming there are ~60 nodes in this Bnet with max number of parents =4; and assuming all nodes are binary, ~ 10^{18} numbers are required for the JPD. How many are required for the Bnet?



iclicker.

A	~10^3
B	~10^{24}
C	~10^4
D	~10^{60}
E	42

Belief network summary

- A belief network is a directed acyclic graph (DAG) that effectively expresses independence assertions among random variables.
- The parents of a node *X* are those variables on which *X* directly depends.
- Consideration of causal dependencies among variables typically help in constructing a Bnet

Next Class

Bayesian Networks Representation

- **Additional Dependencies** encoded by BNets
- More **compact representations** for CPT
- Very simple but extremely useful BNet (**Naïve Bayes Classifier**)