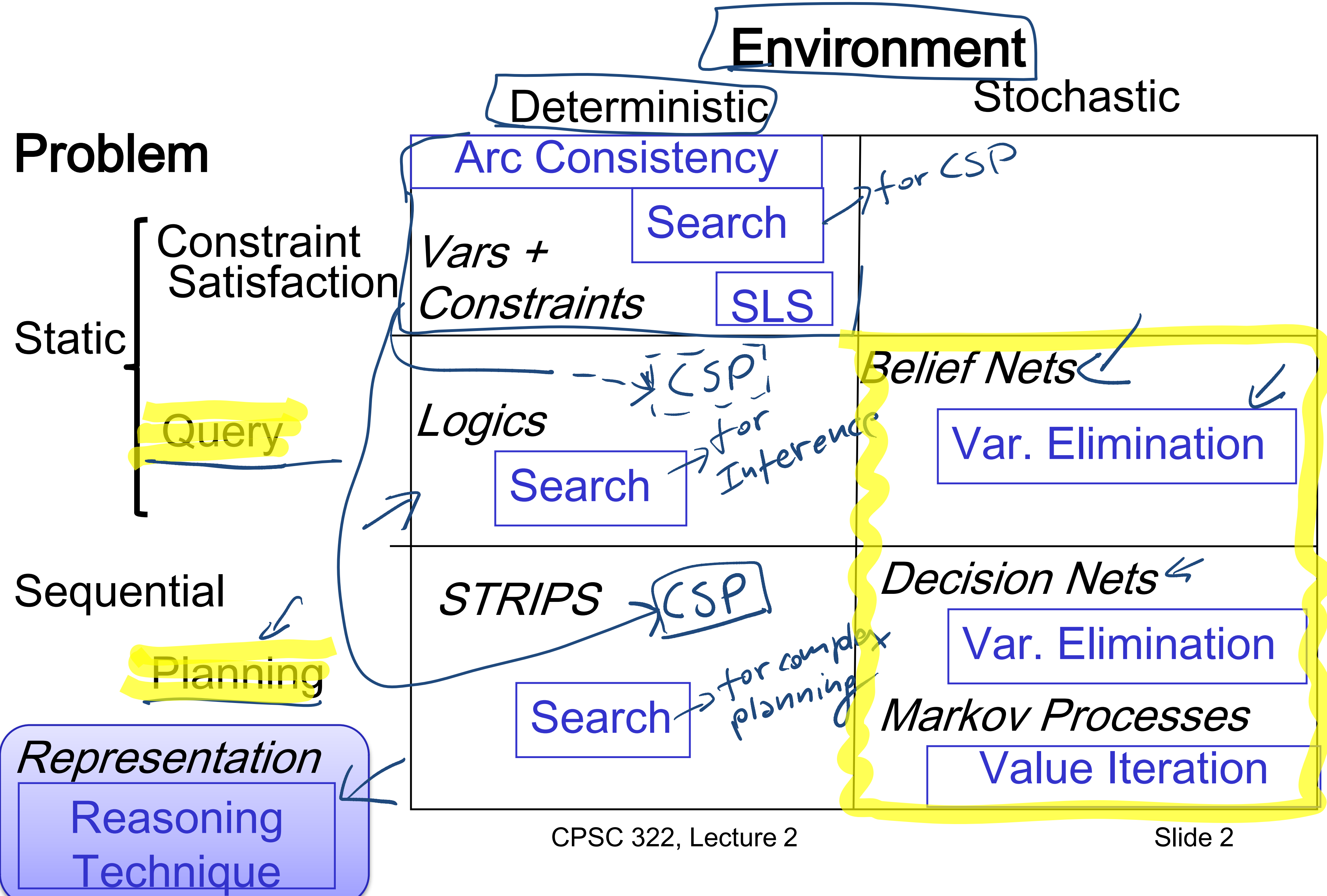
# Reasoning Under Uncertainty: Belief Networks

Computer Science cpsc322, Lecture 27

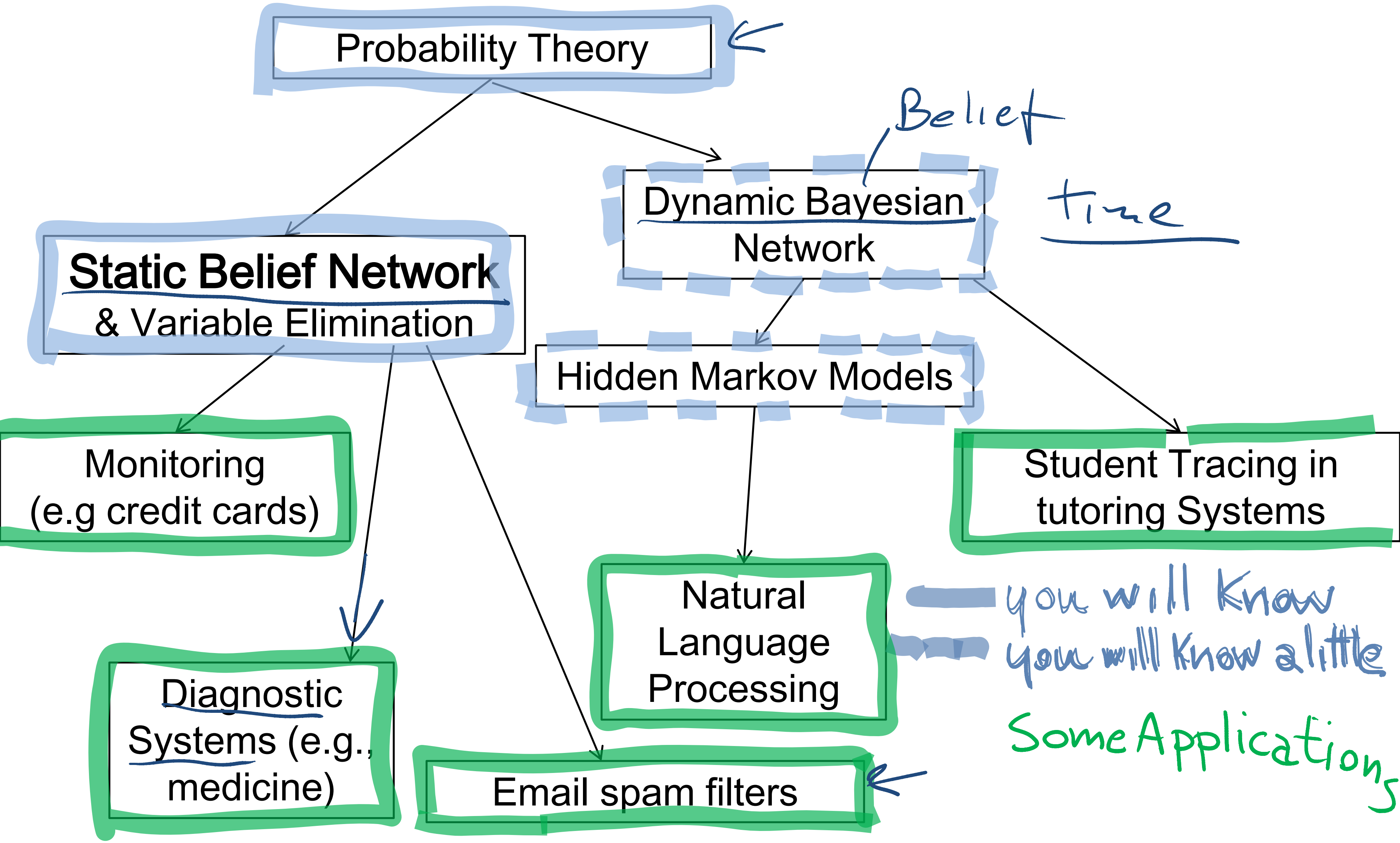
*(Textbook Chpt 6.3)*

March, 16, 2009

# Big Picture: R&R systems



# Answering Query under Uncertainty



# Key points Recap

- We model the environment as a set of …. vars random

$X_1 \ldots X_n$   $P(X_1 \ldots X_n)$

- Why the joint is not an adequate representation ?

"Representation, reasoning and learning" are "exponential" in ….. # vars

**Solution:** Exploit marginal&**conditional** independence

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

But how does independence allow us to simplify the joint? CHAIN RULE!

# 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

B

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

A

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

M J

**Minor earthquakes** may occur and sometimes the set off the alarm.

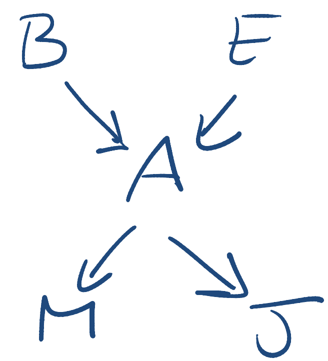
E

**Variables:** B A M J E

**Joint** has $2^5$ 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

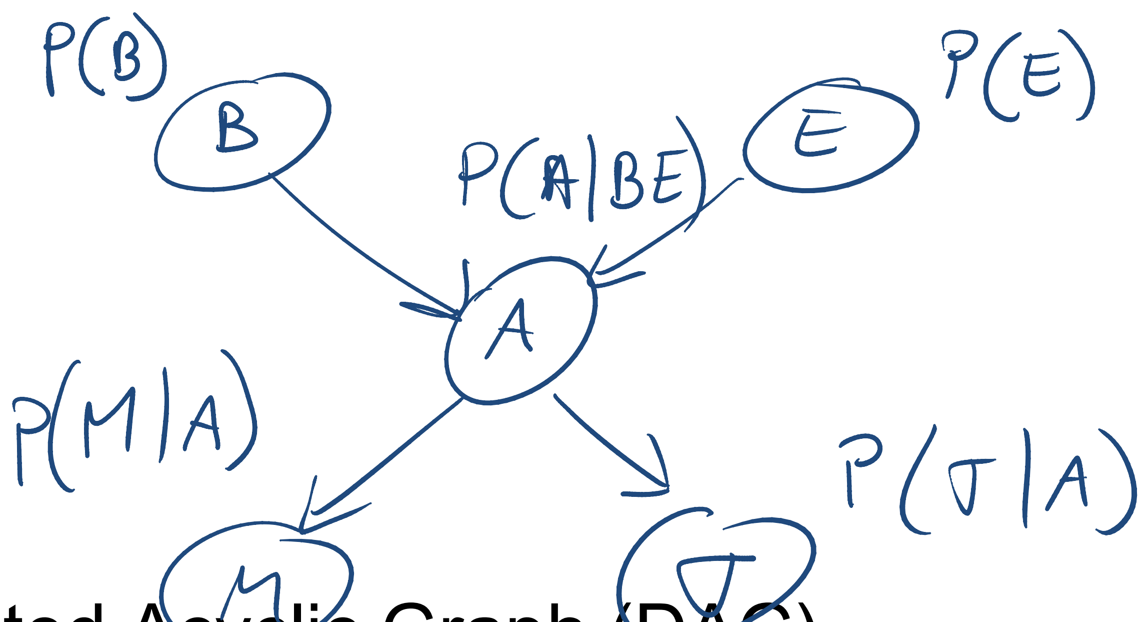
$$= P(B)\, P(E \mid \cancel{B})\, \underline{P(A \mid B, E)}\, P(M \mid \cancel{B}, \cancel{E}, A)\, P(J \mid \cancel{B}, \cancel{E}, A, \cancel{M})$$

- Simplify according to marginal&conditional independence

# Belief Nets: Structure + Probs

$$P(B) * P(E) * P(A|\underline{B,E}) * P(M|\underline{A}) * P(J|\underline{A})$$

- 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)

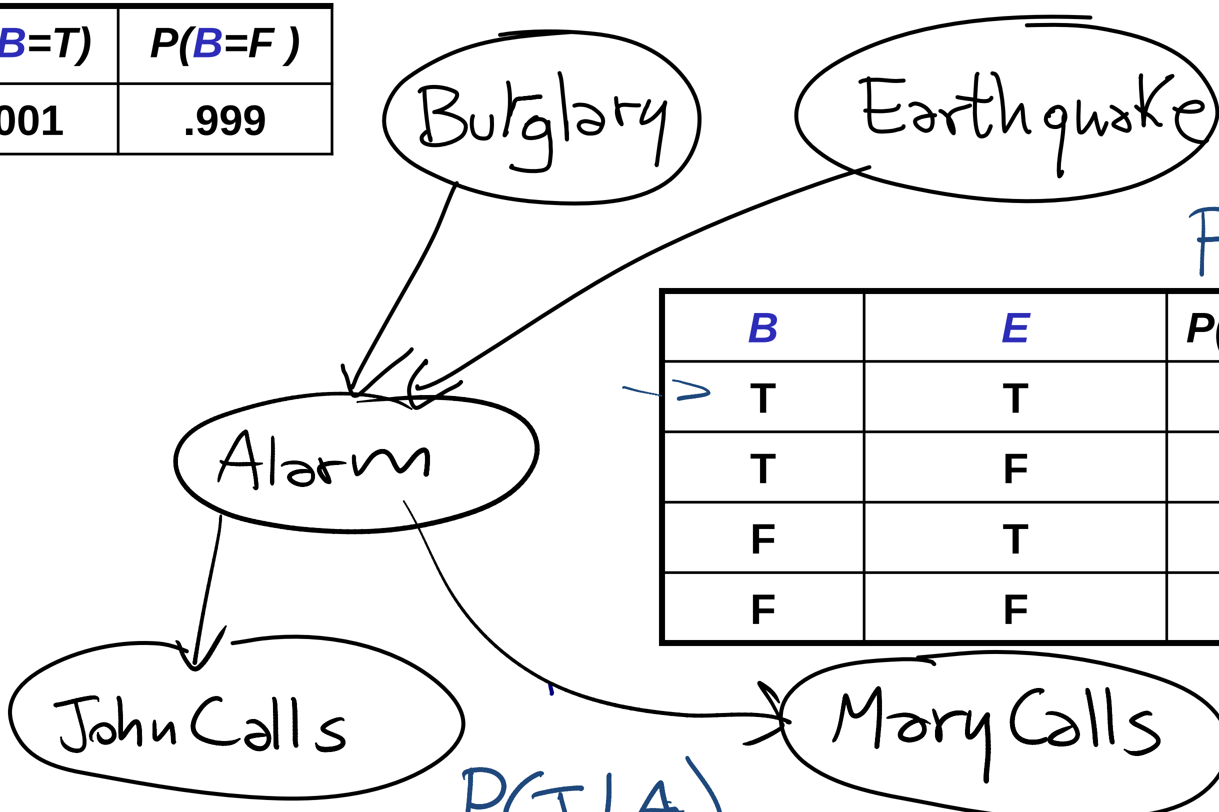
# Burglary: complete BN

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

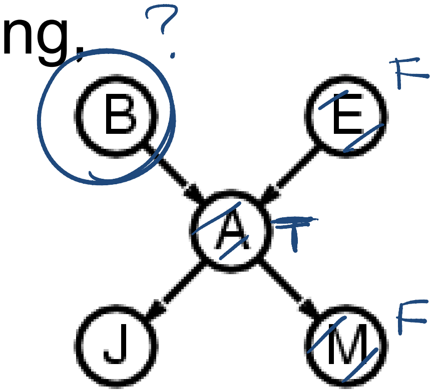
- Belief Networks
  - Build sample BN
  - Intro Inference, Compactness, Semantics
  - More Examples

# Burglary Example: Bnets inference

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

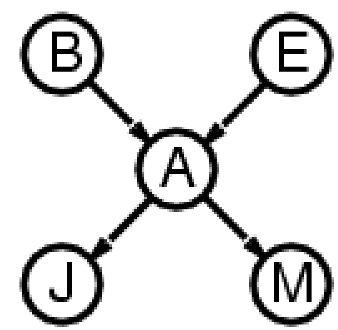
(Ex1) I'm at work,

- neighbor John calls to say my alarm is ringing,
- neighbor Mary doesn't call.
- No news of any earthquakes.
- Is there a burglar?



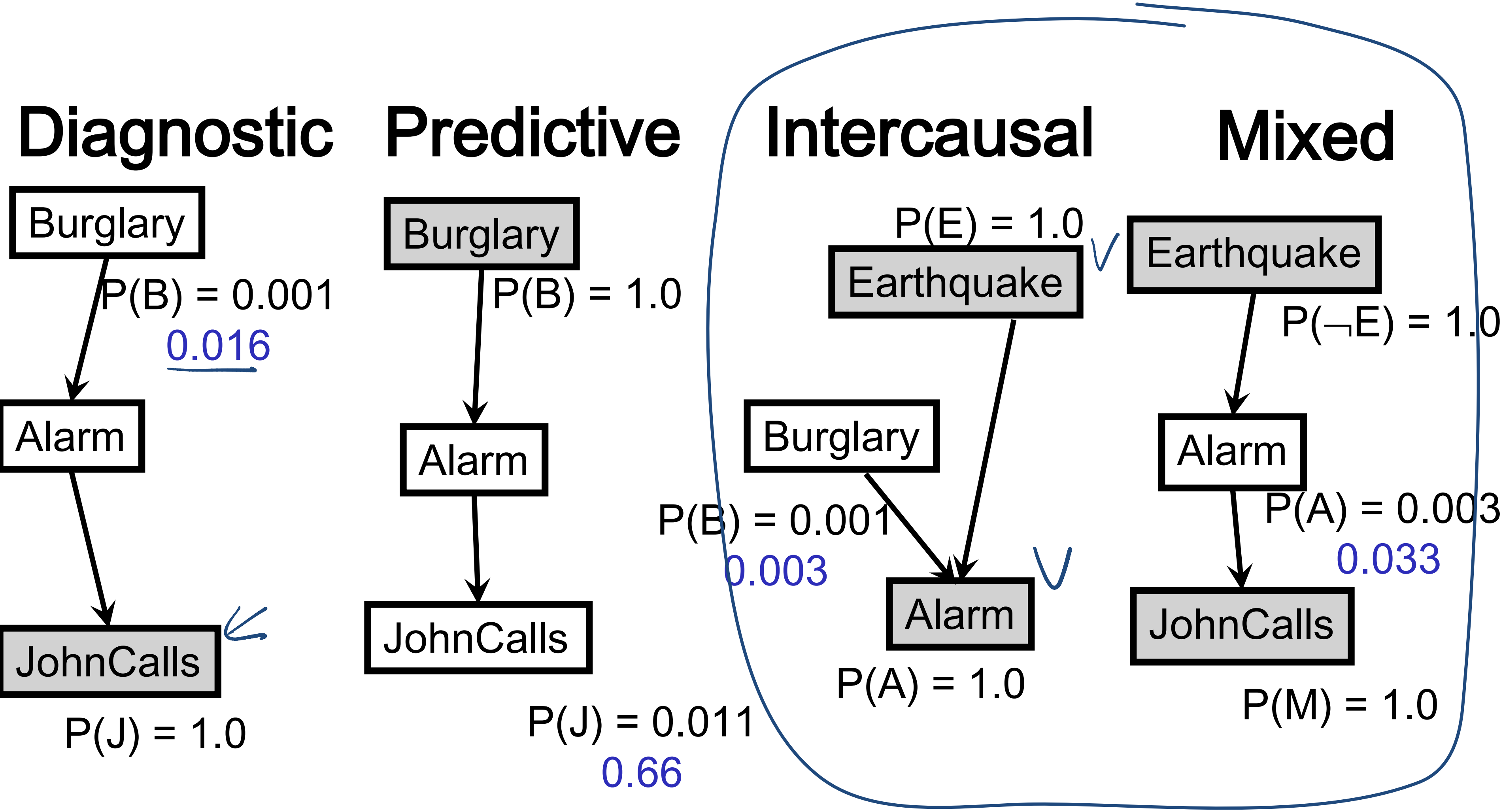

(Ex2) I'm at work,

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

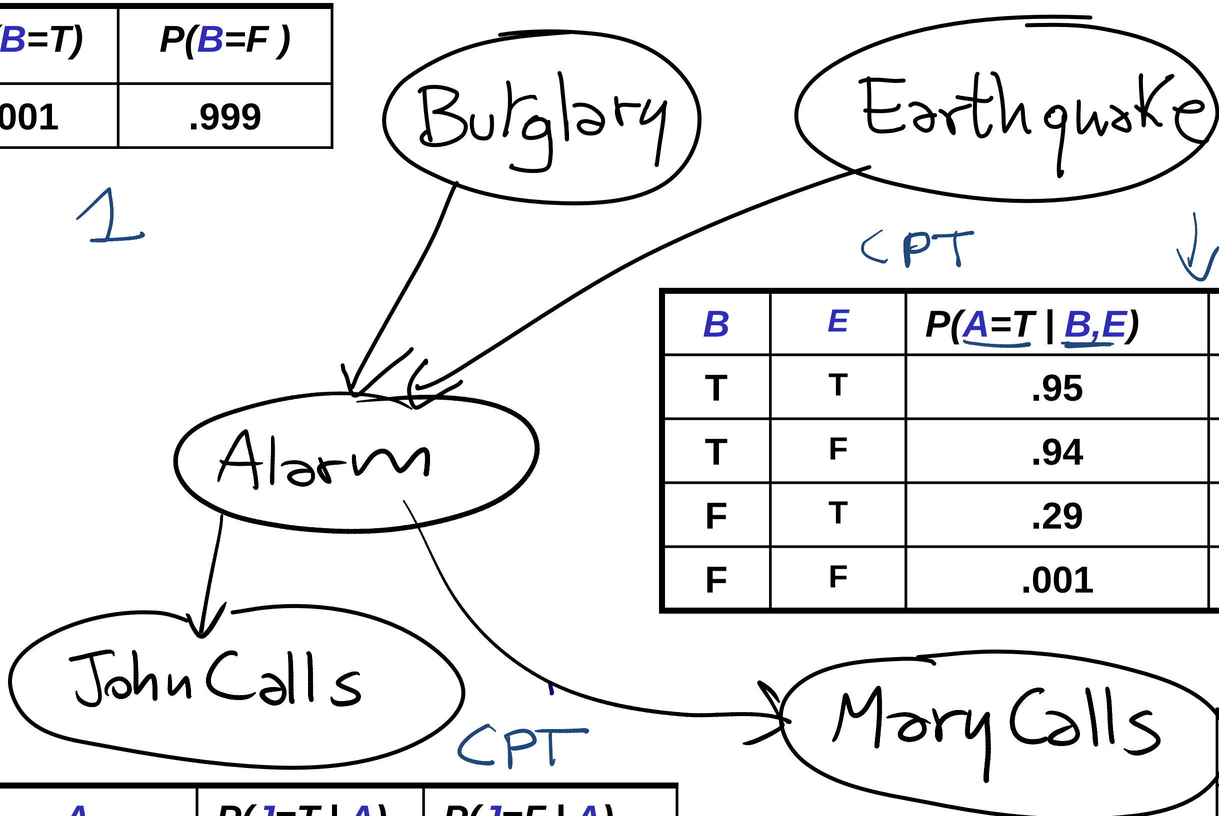



AIspace

# Bayesian Networks – Inference Types



# BNnets: Compactness




CPT

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

1

CPT

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

1

CPT

| 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 |

$2^2$

4

CPT

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

2

CPT

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

2

$1+1+4+2+2=10$ vs $2^5-1=31$

# BNets: Compactness

Conditional Probability Table


$2^k$ possible combinations

## In General:

A CPT for boolean $X_i$ with $k$ boolean parents has $2^k$ rows for the combinations of parent values

each row is a distrib. for $X_i$

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 $O(n\,2^k)$ numbers

$2^n$

For $k << n$, 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:

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

Simplify according to marginal&conditional independence

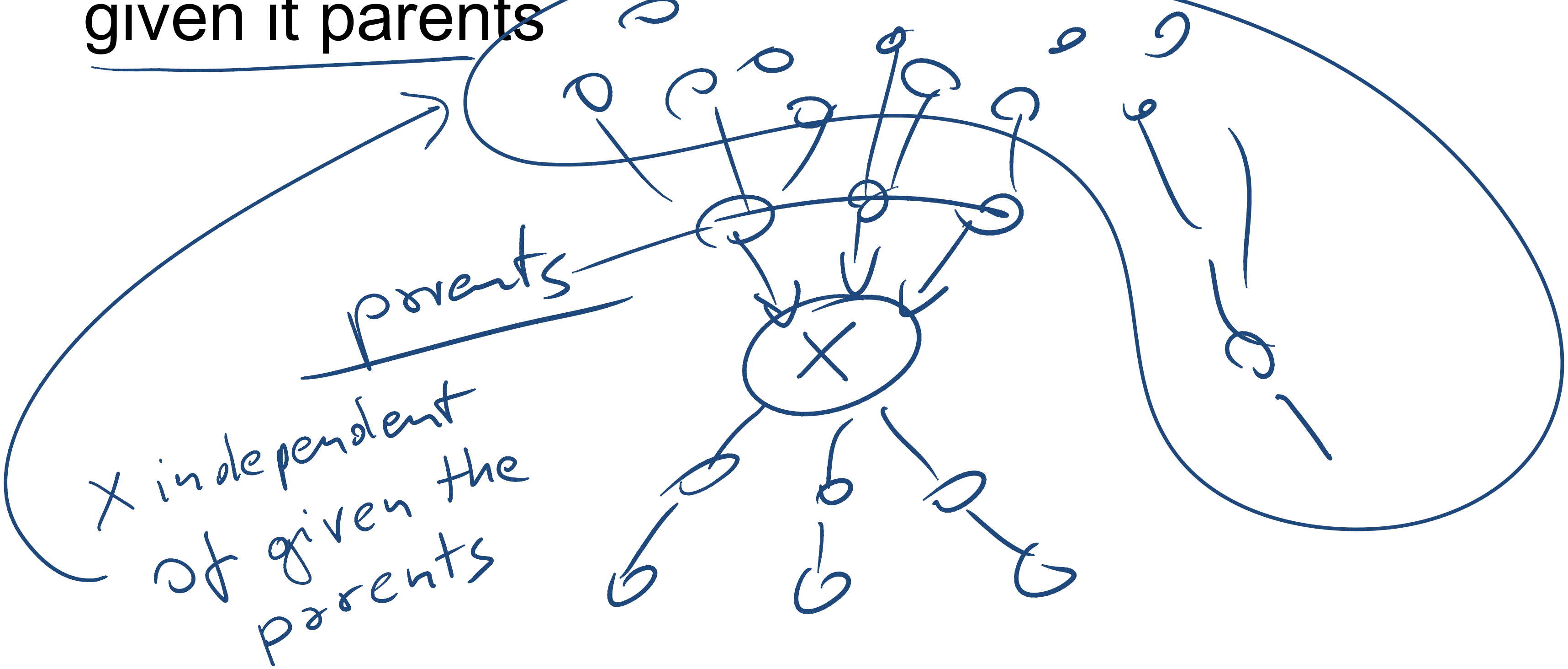
- Express remaining dependencies as a network
  - Each var is a node
  - For each var, the conditioning vars are its parents
  - Associate to each node corresponding conditional probabilities

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

# BNets: Construction General Semantics (cont')

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

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



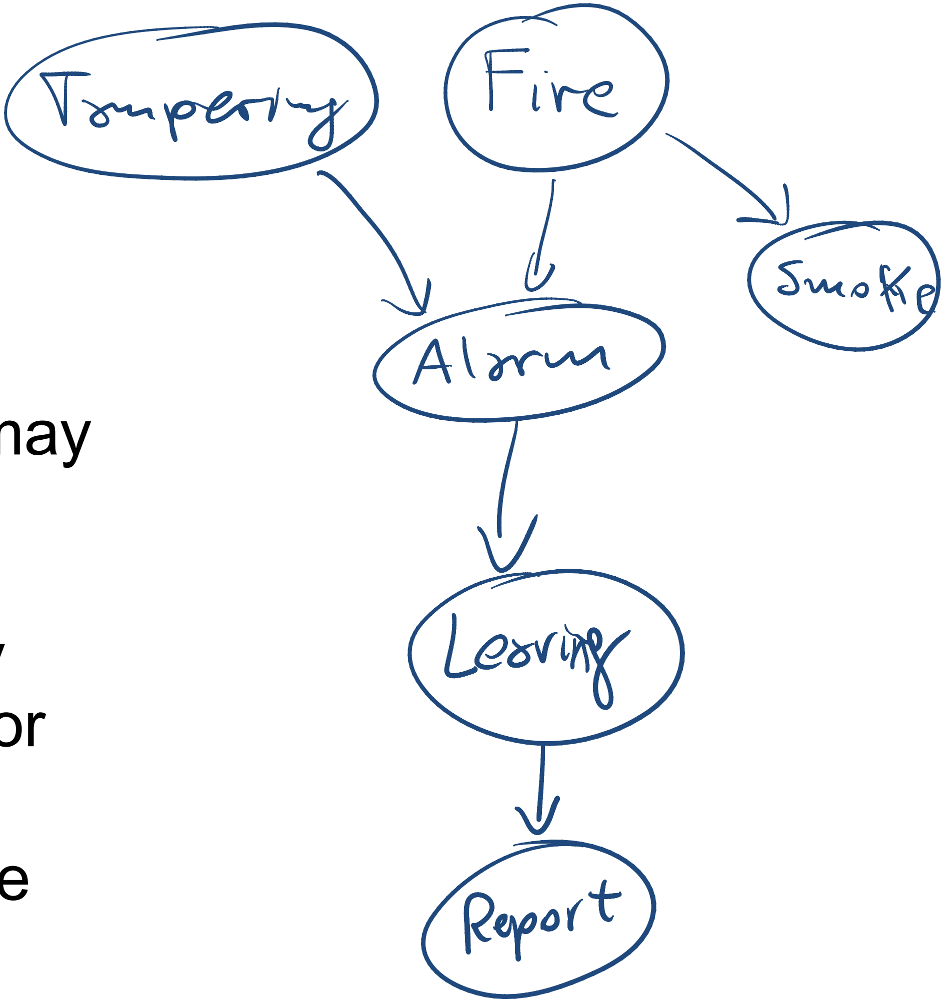
# Lecture Overview

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

# Other Examples: Fire Diagnosis (textbook Ex. 6.10)

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.



# 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)

- **Patient's aching hands** example (ex. 6.14)

- Several other examples on AIspace

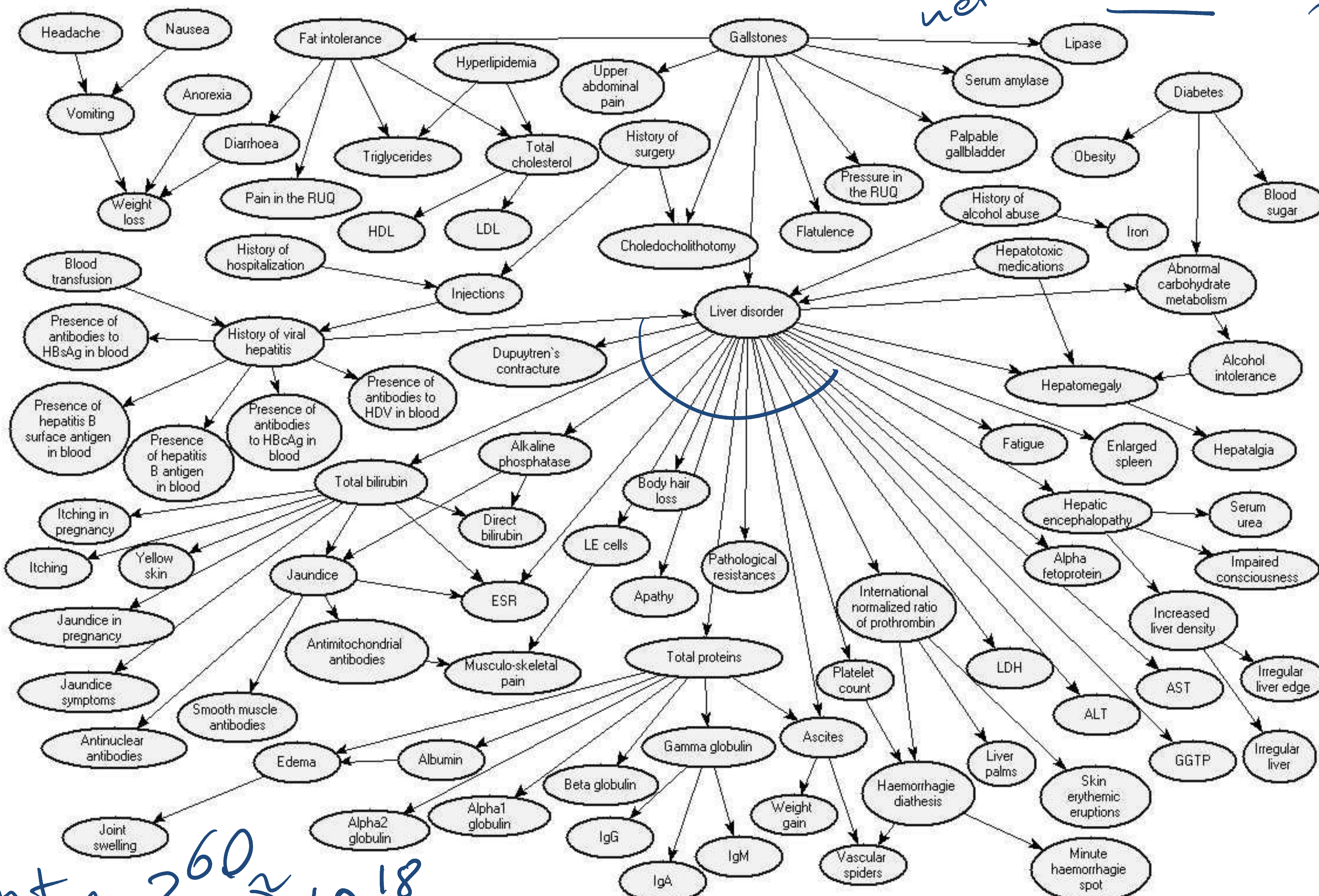
# Realistic BNet: Liver Diagnosis

Source: Onisko et al., 1999

this network

60 2 4

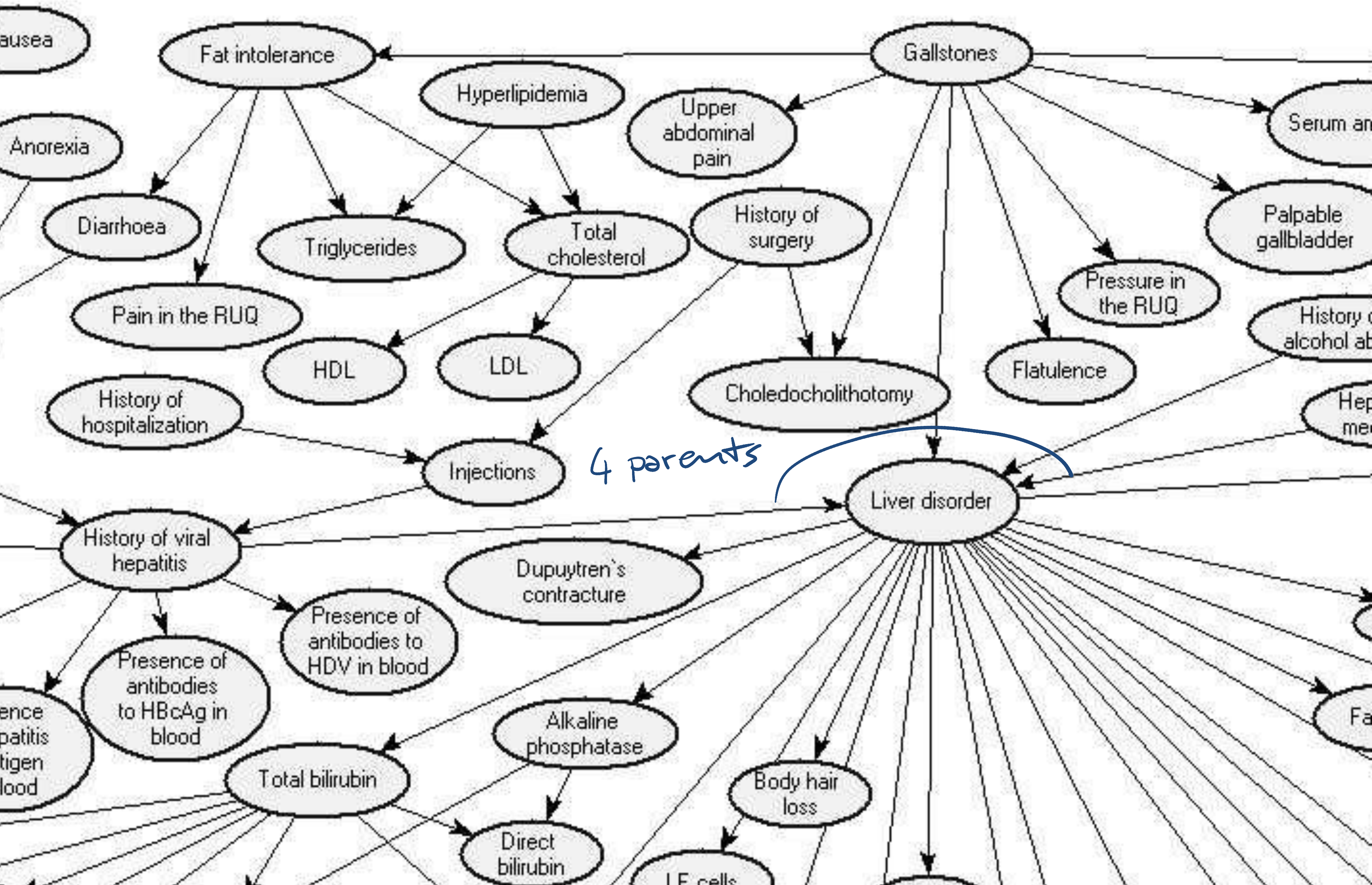
$\sim 10^3$




Joint $\sim 2^{60} \cong 10^{18}$

# Realistic BNet: Liver Diagnosis

Source: Onisko et al., 1999



# 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

# Next Class

Bayesian Networks Representation

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

# 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