CS 730/730W/830: Intro AI

1 handout: slides
final blog entries were due

Bayesian Networks

The Alarm Domain

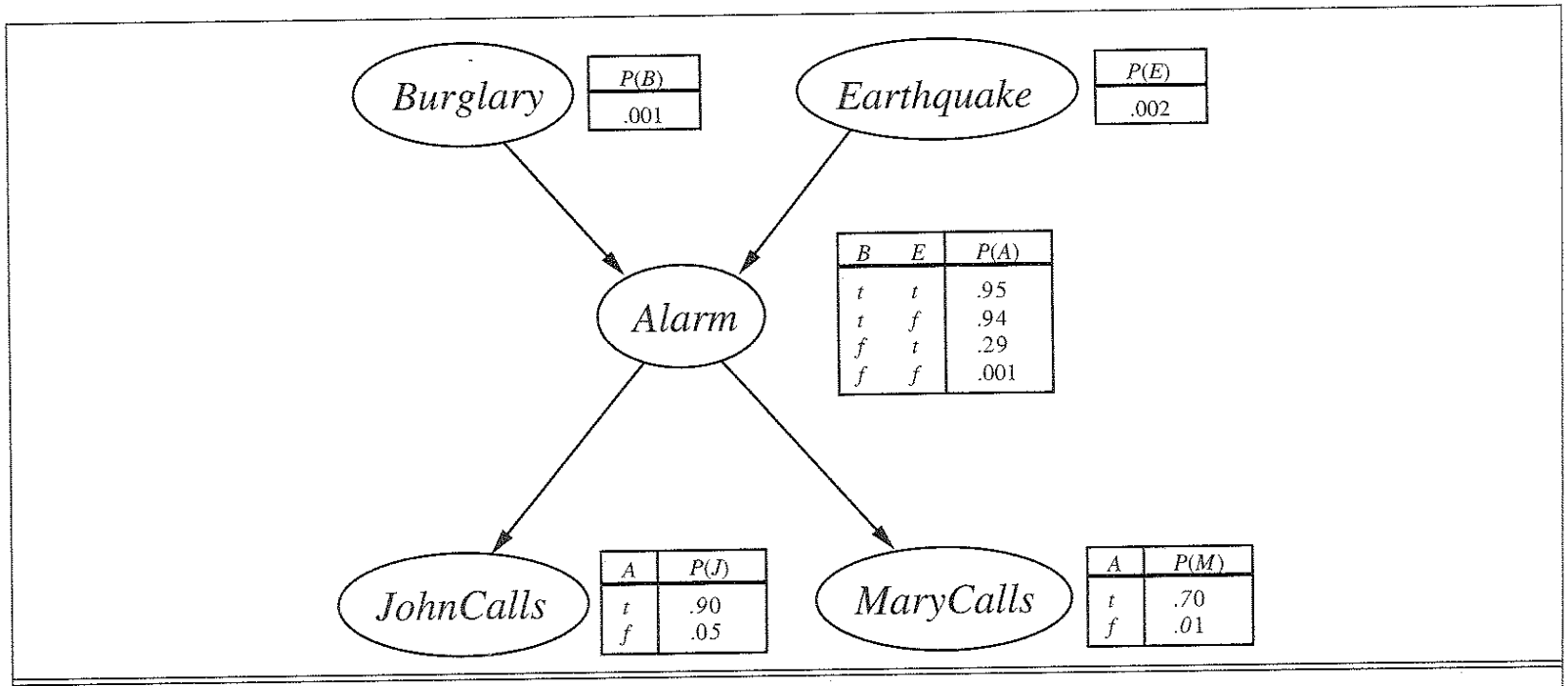

Figure 14.2 A typical Bayesian network, showing both the topology and the conditional probability tables (CPTs). In the CPTs, the letters B, E, A, J, and M stand for *Burglary*, *Earthquake*, *Alarm*, *JohnCalls*, and *MaryCalls*, respectively.

Bayes Nets Reminder

In general:

$$P(x_1, \ldots, x_n) = P(x_n | x_{n-1}, \ldots, x_1) P(x_{n-1}, \ldots, x_1)$$

Bayes Nets Reminder

In general:

$$\begin{aligned} P(x_1, \ldots, x_n) &= P(x_n | x_{n-1}, \ldots, x_1) P(x_{n-1}, \ldots, x_1) \\ &= \prod_{i=1}^{n} P(x_i | x_{i-1}, \ldots, x_1) \end{aligned}$$

Bayes Net specifies independence:

$$P(X_i | X_{i-1}, \ldots, X_1) = P(X_i | parents(X_i))$$

joint distribution:

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i))$$

What is distribution of X given evidence e and unobserved Y?

Approximate Inference

Sampling According to the Joint Distribution

sample values for variables, working top down

directly implements the semantics of the network
'generative model'

each sample is linear time

Rejection Sampling

What is distribution of X given evidence e and unobserved Y?

Draw worlds from the joint, rejecting those that do not match e.
Look at distribution of X.

each sample is linear time, but overall slow if e is unlikely

Likelihood Weighting

What is distribution of X given evidence e and unobserved Y?

ChooseSample (e)

$w \leftarrow 1$

for each variable V_i in topological order:

  if $(V_i = v_i) \in e$ then

    $w \leftarrow w \cdot P(v_i | parents(v_i))$

  else

    $v_i \leftarrow$ sample from $P(V_i | parents(V_i))$

(afterwards, normalize samples so all w's sum to 1)

uses all samples, but needs lots of samples if e are late in ordering

Break

- exam 3: calculator, review session May 4
- projects

Exact Inference in Bayesian Networks

Enumeration Over the Joint Distribution

What is distribution of X given evidence e and unobserved Y?

$$\begin{aligned} P(X|e) &= \frac{P(e|X)P(X)}{P(e)} \\ &= \alpha P(X,e) \\ &= \alpha \sum_y P(X,e,y) \\ &= \alpha \sum_y \prod_{i=1}^{n} P(V_i|parents(V_i)) \end{aligned}$$

Example

$$
\begin{aligned}
P(B|j,m) &= \frac{P(j,m|B)P(B)}{P(j,m)} \\
&= \alpha P(B,j,m) \\
&= \alpha \sum_e \sum_a P(B,e,a,j,m) \\
&= \alpha \sum_e \sum_a \prod_{i=1}^{n} P(V_i|parents(V_i)) \\
P(b|j,m) &= \alpha \sum_e \sum_a P(b)P(e)P(a|b,e)P(j|a)P(m|a) \\
&= \alpha P(b) \sum_e P(e) \sum_a P(a|b,e)P(j|a)P(m|a)
\end{aligned}
$$

[draw tree]

Variable Elimination

$$P(B|j,m) \quad = \quad \alpha P(B) \sum_e P(e) \sum_a P(a|B,e) P(j|a) P(m|a)$$

factors = tables = $f_{varsused}(dimensions)$.
eg: $f_A(A,B,E)$, $f_M(A)$
multiplying factors: table with union of variables
summing reduces table

Variable Elimination

eliminating variables: eg $P(J|b)$

$$P(J|b) = \alpha P(b) \sum_e P(e) \sum_a P(a|b,e) P(J|a) \sum_m P(m|a)$$

all vars not ancestor of query or evidence are irrelevant!

EOLQs

- What question didn't you get to ask today?
- What's still confusing?
- What would you like to hear more about?

Please write down your most pressing question about AI and put it in the box on your way out.
Thanks!