Artificial Intelligence-BSCE-306L

Module 5

Uncertain Knowledge and Reasoning

Dr . Saurabh Agrawal

Faculty Id: 20165

School of Computer Science and Engineering

VIT, Vellore-632014

Tamil Nadu, India

Outline

- ❑Quantifying Uncertainty
- ❑Bayes Rule
- ❑Bayesian Belief Network
- ❑Approximate Inference in Bayesian networks

Uncertainty

- When an agent knows enough facts about its environment, the logical plans and actions produces a guaranteed work.
- Unfortunately, *agents never have access to the whole truth about their environment*. Agents act under **uncertainty.**

Nature of Uncertain Knowledge

- The Diagnosis : medicine, automobile repair, or whatever-is a task that almost always involves uncertainty
- Let us try to write rules for dental diagnosis using first-order logic, so that we can see how the logical approach breaks down. Consider the following rule:
  - *∀p Symptom(p, Toothache) ⇒ Disease(p, Cavity) .*
- The problem is that this rule is wrong.
- Not all patients with toothaches have cavities; some of them have gum disease, swelling, or one of several other problems
  - *∀p Symptom(p, Toothache) ⇒ Disease(p, Cavity) V Disease(p, GumDisease) V Disease(p, Swelling) . . .*
- to make the rule true, we have to add almost unlimited list of possible causes.
- We could try a causal rule:
  - *∀p Disease(p, Cavity) ⇒ Symptom(p, Toothache).*
- But this rule is also not right either; not all cavities cause pain
- Toothache and a Cavity are unconnected, so the judgement may go wrong.

Nature of Uncertain Knowledge...

- This is a type of the medical domain, as well as most other judgmental domains: law, business, design, automobile repair, gardening, dating, and so on.
- The agent take action, only a **degree of belief** in the relevant sentences.
- Our main tool for dealing with degrees of belief will be **probability theory**
- **The Probability** assigns to each sentence a numerical degree of belief between 0 and 1.

Probability

- Probabilities are used to compute the truth of given statement, written as numbers between 0 and 1, that describes how likely an event is to occur.
- 0 indicates impossibility and 1 indicates certainly.
  - 1. Tossing a coin 2. Tolling a dice
- Probability based reasoning
  - understanding from knowledge
  - how much of uncertainty present in that event.

Quantifying Uncertainty

Probability

- *Probability provides a way of* **summarizing** *the uncertainty, that comes from our laziness and ignorance.*
- Toothache problem - an 80% chance , a probability of 0.8 that the patient has a cavity if he or she has a toothache.
- The 80% summarizes those cases, but both toothache and cavity are unconnected.
- The missing 20% summarizes, all other possible causes of toothache, that we are too lazy or ignorant to confirm or deny.

Quantifying Uncertainty

- Probabilities between 0 and **1** correspond to intermediate degrees of belief in the truth of the sentence.
- The sentence itself is in *fact* either true or false.
- It is important to note that a degree of belief is different from a degree of truth.
- A probability of 0.8 does not mean "80% true" but rather an 80% degree of belief-that is, a fairly strong expectation.
- Thus, probability theory makes the same ontological commitment as logic-namely, that facts either do or do not hold in the world.
- Degree of truth, as opposed to degree of belief, is the subject of **fuzzy logic**

Quantifying Uncertainty

- In probability theory, a sentence such as
- "The probability that the patient has a cavity is 0.8“,
- is about the agent's beliefs, not directly about the world.
- These percepts create the **evidence,** which are based on probability statements.
- All probability statements must indicate the evidence with respect to that probability is being assessed.
- If an agent receives new percepts, its probability assessments are updated to reflect the new evidence.

Random Variable

- Referring to a "part" of the world, whose "status" is initially unknown
- We will use lowercase for the names of values
  - $P(a) = 1 - P(\neg a)$
- Tossing coin : $P(h) = 1 - P(\neg h)$: $(0.5 = 1 - 0.5)$
- Rolling dice : $P(n) = 1 – P(\neg n)$: $(0.16 = 1 - 0.84)$

Types of random variables

- Boolean random variables
  - Cavity domain (true,false), if Cavity = true then cavity, or
  - if Cavity = false then ¬cavity
- Discrete random variables – countable domain
  - *Weather* might be *(sunny, rainy, cloudy, snow)*
- Continuous random variables – finite set real numbers with equal intervals e.g. interval(0.1)

Atomic events

- The concept of an **atomic event** is useful in understanding the foundations of probability theory.
- It is a complete specification of the state of the world about which the agent is uncertain.
- It can be an assignment of particular values, to all the variables of which the world is composed

Atomic events...

- Atomic events have some important properties
- They are mutually exclusive -at most one can actually be the case.
- The set of all possible atomic events is exhaustive – at least one must be the case.
- Any particular atomic event entails the truth or falsehood of every proposition, whether simple or complex
- Any proposition is logically equivalent to the disjunction of all atomic events that required the truth of proposition.

Prior Probability

- The **unconditional** or **prior probability** associated with a proposition *a*, is the degree of belief according to the absence of any other information;
- it is written as *P(a)*.
- For example, if the prior probability that one have a cavity is 0.1, then we would write
  - *P(Cavity = true)* = *0.1* or *P(cavity)* = 0.1 .
- It is important to remember that *P(a)* can be used only when there is no other information.

Prior Probability...

- we will use an expression **P(** *Weather),* which denotes a ***vector*** of values, for the probabilities of each individual state of the weather.
  - *P(Weather = sunny) = 0.7*
  - *P(Weather = rain) = 0.2*
  - *P(Weather = cloudy) = 0.08*
  - *P(Weather = snow) = 0.02 .*
- we may simply write
- *P(Weather) = (0.7, 0.2, 0.08, 0.02) .*
- This statement defines a prior **probability distribution** for the random variable *Weather.*

Quantifying Uncertainty

Conditional Probability

- The **conditional** or **posterior** probabilities notation is $P(a|b)$,
- where a and b are any proposition.
- This is read as "the probability of a, given that *all* we know is b."
- For example, $P(cavity|\ toothache) = 0.8$
- if a patient is observed to have a toothache and no other information is yet available, then the probability of the patient's having a cavity will be 0.8.

Conditional probabilities...

- Conditional probabilities can be defined in terms of unconditional probabilities.
- The equation is

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}$$

- whenever $P(b) > 0$.
- This equation can also be written as
- $P(a \wedge b) = P(a/ b)\, P(b)$ which is called the **product rule.**

Basic Axioms of Probability

- All probabilities are between 0 and 1. For any proposition a,

$$0 \leq P(a) \leq 1$$

- Necessarily true (i.e., valid) propositions have probability I, and necessarily false (i.e., unsatisfiable) propositions have probability 0.

$$P(true) = 1 \qquad P(false) = 0 .$$

- The probability of a disjunction is given by

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b) .$$

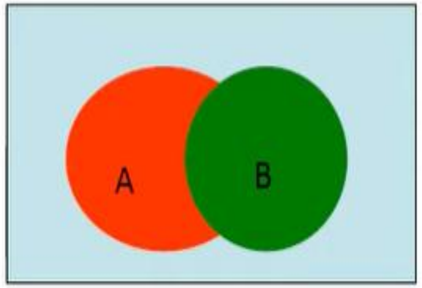
17

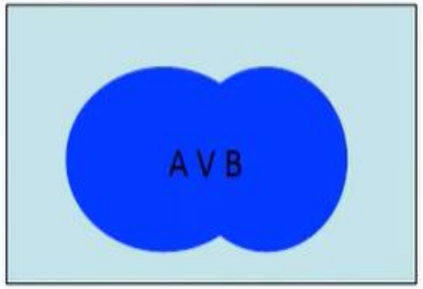
❑Quantifying

Axioms of probability

- $0 \leq P(A) \leq 1$

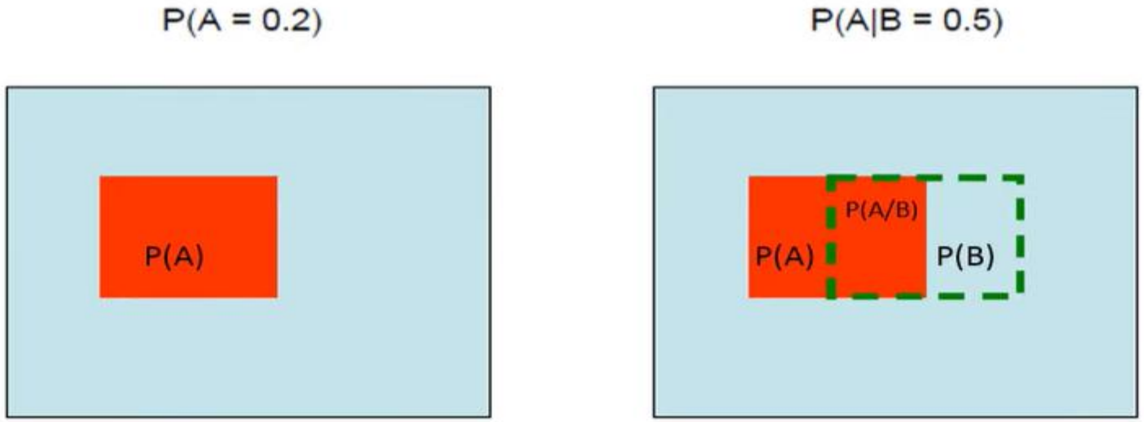
- 2. $P(\mathit{true}) = 1$, $P(\mathit{false}) = 0$
- 3. $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$



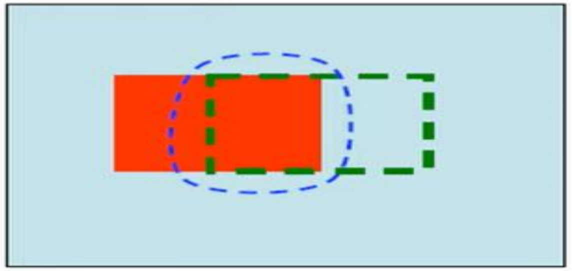


Conditional Probability...

- P(A = 1 | B = 1):
- The fraction of cases where A is true if B is true



- P(A,B) = P(A | B)*P(B)
- this is one of the most powerful rules in probabilistic reasoning



Using the Axioms

- How can we use the axioms to prove that:
  - $P(A) = 1 - P(\neg A)$
- Prior Probability – Degree-of-belief in an event, in the absence of any other information
  - P(rain tomorrow) = 0.8
  - P(no-rain tomorrow) = 0.2

Conditional Probability

- What is the probability of an event, given knowledge of another event
- Example:
  - P(raining | sunny)
  - P(raining | cloudy)
  - P(raining | cloudy, cold)

Conditional probability…

- In some cases, given knowledge of one or more random variables, we can improve upon our prior belief of another random variable
- For example:
  - p(slept in movie) = 0.5
  - p(slept in movie | liked movie) = 1/3
  - p(didn't sleep in movie | liked movie) = 2/3

Liked movie	Slept	P
1	1	0.2
1	0	0.4
0	0	0.1
0	1	0.3

Bayes Rule

- Bayes' theorem is also known as **Bayes' rule, Bayes' law**, or **Bayesian reasoning**, which determines the probability of an event with uncertain knowledge.
- In probability theory, it relates the conditional probability and marginal probabilities of two random events.
- Bayes' theorem was named after the British mathematician **Thomas Bayes**.
- The **Bayesian inference** is an application of Bayes' theorem, which is fundamental to Bayesian statistics.
- **It is a way to calculate the value of P(B|A) with the knowledge of P(A|B).**
- Bayes' theorem allows updating the probability prediction of an event by observing new information of the real world.

Bayes Rule

- **Example:** If cancer corresponds to one's age then by using Bayes' theorem, we can determine the probability of cancer more accurately with the help of age.
- Bayes' theorem can be derived using product rule and conditional probability of event A with known event B:
- **As from product rule we can write:** **$P(A \wedge B) = P(A|B)\, P(B)$**
- or
- Similarly, the probability of event B with known event A: **$P(A \wedge B) = P(B|A)\, P(A)$**
- Equating right hand side of both the equations, we will get:

$$\boxed{P(A|B) = \frac{P(B|A)\, P(A)}{P(B)}} \quad \text{....(a)}$$

Bayes Rule

- The above equation (a) is called as **Bayes' rule** or **Bayes' theorem**. This equation is basic of most modern AI systems for **probabilistic inference**. It shows the simple relationship between joint and conditional probabilities.
- Here, **P(A|B)** is known as **posterior**, which we need to calculate, and it will be read as Probability of hypothesis A when we have occurred an evidence B.
- **P(B|A)** is called the likelihood, in which we consider that hypothesis is true, then we calculate the probability of evidence.
- **P(A)** is called the **prior probability**, probability of hypothesis before considering the evidence
- **P(B)** is called **marginal probability**, pure probability of an evidence.

Bayes Rule

- In the equation (a), in general, we can write **P (B) = P(A)*P(B|Ai)**, hence the Bayes' rule can be written as:

$$P(A_i|B) = \frac{P(A_i) * P(B|A_i)}{\sum_{i=1}^{k} P(A_i) * P(B|A_i)}$$

- Where A_1, A_2, A_3,........, A_n is a set of mutually exclusive and exhaustive events.

Bayes Rule

❑Applying Bayes' rule:

❑Bayes' rule allows us to compute the single term **P(B|A)** in terms of **P(A|B), P(*B*),** and **P(A)**.

❑This is very useful in cases where we have a good probability of these three terms and want to determine the fourth one.

❑Suppose we want to perceive the effect of some unknown cause, and want to compute that cause, then the Bayes' rule becomes:

$$P(\text{cause} \mid \text{effect}) = \frac{P(\text{effect} \mid \text{cause})\, P(\text{cause})}{P(\text{effect})}$$

Bayes Rule

- **Example-1:Question: From a standard deck of playing cards, a single card is drawn. The probability that the card is king is 4/52, then calculate posterior probability P(King|Face), which means the drawn face card is a king card.**

- **Solution:**

$$P(\text{king}|\text{face}) = \frac{P(\text{Face}|\text{king}) * P(\text{King})}{P(\text{Face})} \quad ....(i)$$

- **P(king): probability that the card is King= 4/52= 1/13**
- **P(face): probability that a card is a face card= 3/13**
- **P(Face|King): probability of face card when we assume it is a king = 1**
- **Putting all values in equation (i) we will get:**

$$P(\text{king}|\text{face}) = \frac{1 * (\frac{1}{13})}{(\frac{3}{13})} = \mathbf{1/3}, \text{ it is a probability that a face card is a king card.}$$

Bayes Rule

Example – Hypothesis for Flu based on Symptoms

- Given :
  - P(A) = Symptom of Flu = 0.00001
  - P(B/A) = Probability of Symptoms gives Flu = 0.95
  - P(B) = Your symptom of flu = 0.01 (head ache or running nose, 1 in 100)

- Compute P(A/B)

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

  - P(A/B) = 0.95 * 0.00001 / 0.01
  - = 0.00095 (less than one in thousand)

- **Important to note the basic mistake**
  - P(A/B) is not equal to P(B/A)
  - 0.00095 is not equal to 0.95

Joint probability distribution

- The **full joint probability distribution** specifies the probability of values to random variables.
- It is usually too large to create or use in its explicit form.
- Joint probability distribution of two variables X and Y are

Joint probabilities	X	X'
Y	0.20	0.12
Y'	0.65	0.03

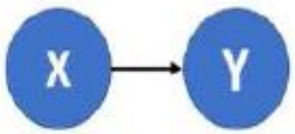
- Joint probability distribution for n variables require 2^n entries with all possible combination.

Drawbacks of joint Probability Distribution

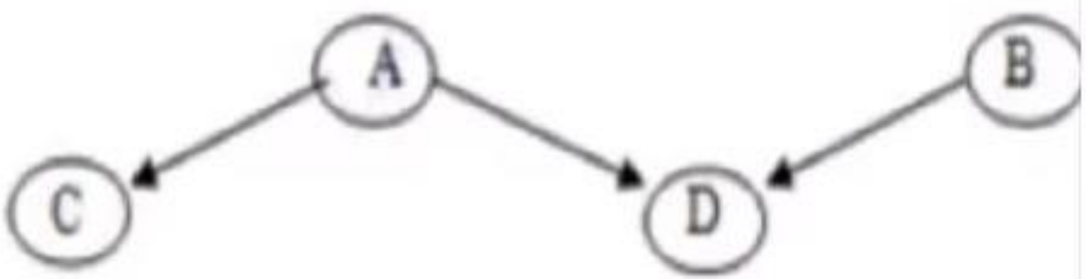
- Large number of variables and grows rapidly
- Time and space complexity are huge
- Statistical estimation with probability is difficult
- Human tends signal out few propositions
- The alternative to this is Bayesian Networks.

Bayesian Networks

- Bayesian Network is to represent the dependencies among variables and to give a brief specification of any full joint probability distribution.
- A Bayesian network is a directed graph in which each nodes are variables and edges are relations.
- The full specification is as follows:
  - 1. A set of random variables makes up the nodes of the network. Variables may be discrete or continuous.
  - 2. A set of directed links or arrows connects pairs of nodes. If there is an arrow from node *X* to node *Y*, X is said to be a parent of *Y*.
  - 3. Each node *X*, has a conditional probability distribution *P(X,(Parents(X,))* that quantifies the effect of the parents on the node. (X is parent of Y)
  - 4. The graph has no directed cycles (and hence is a directed, acyclic graph, or DAG).



Example

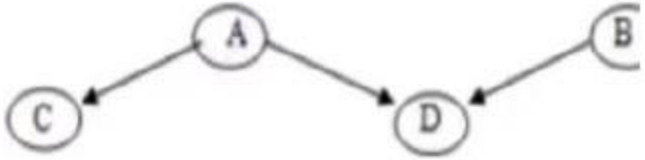


Bayesian Belief Network

- A & B are unconditional, independent, evidence and parent nodes
- C & D are conditional, dependent, hypothesis and child nodes.

Bayesian Belief Network

Conditional Probability Table



P(A)	=	0.3
P(B)	=	0.6
P(C|A)	=	0.4
P(C|~A)	=	0.2
P(D|A, B)	=	0.7
P(D|A, ~B)	=	0.4
P(D|~A, B)	=	0.2
P(D|~A, ~B)	=	0.01

Conditional Probability Tables

P(A)	P(B)
0.3	0.6

A	P(C)
T	0.4
F	0.2

A	B	P(D)
T	T	0.7
T	F	0.4
F	T	0.2
F	F	0.01

- P(A,B,C,D) = P(D|A,B)*P(C|A)*P(B)*P(A)

Example 2

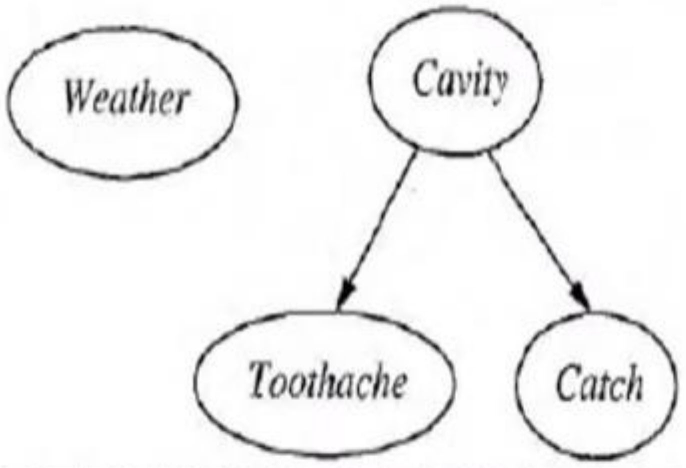


Figure 14.1 A simple Bayesian network in which *Weather* is independent of the other three variables and *Toothache* and *Catch* are conditionally independent, given **Cavity**.

Example -3 - Burglar alarm

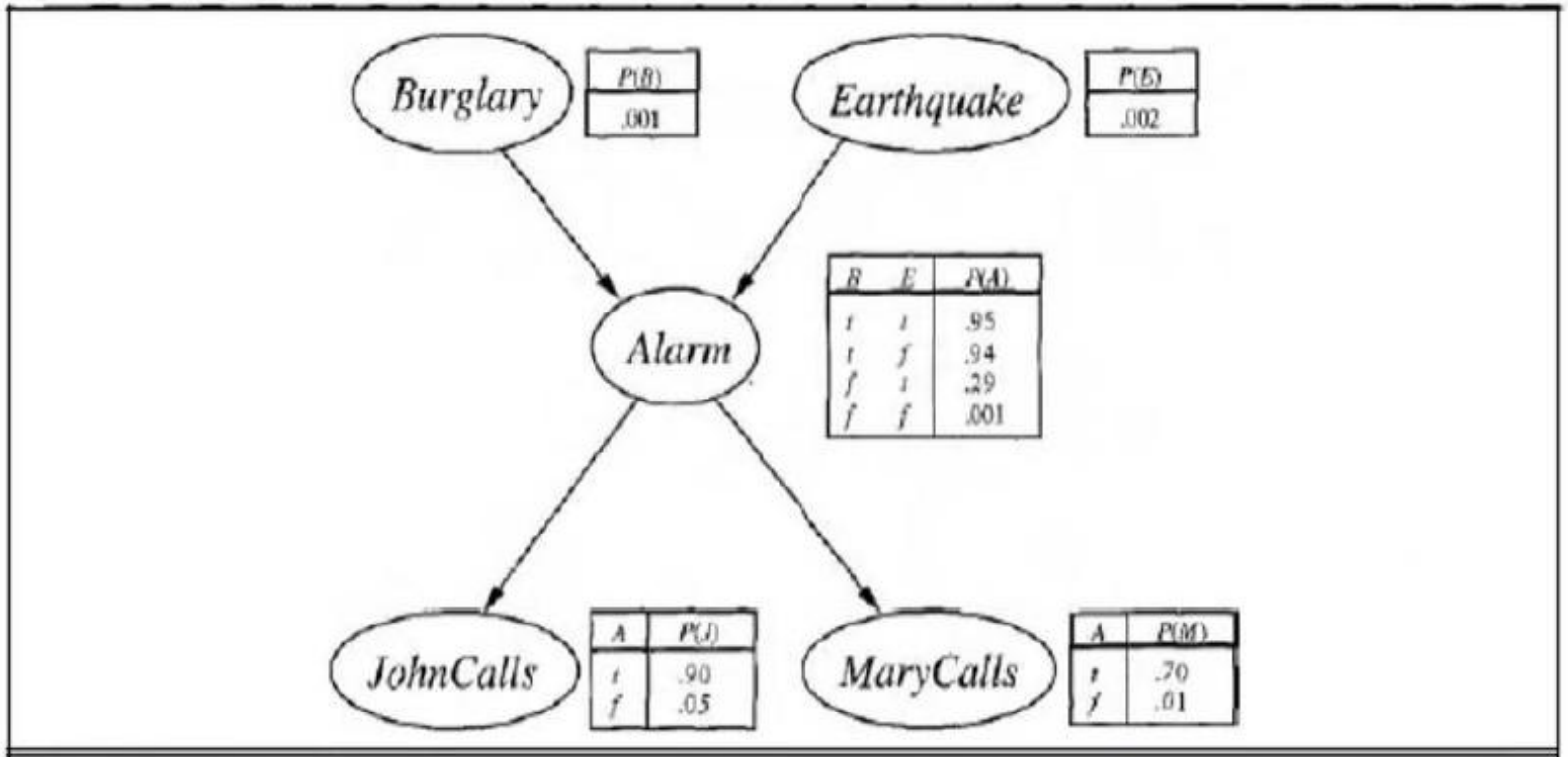


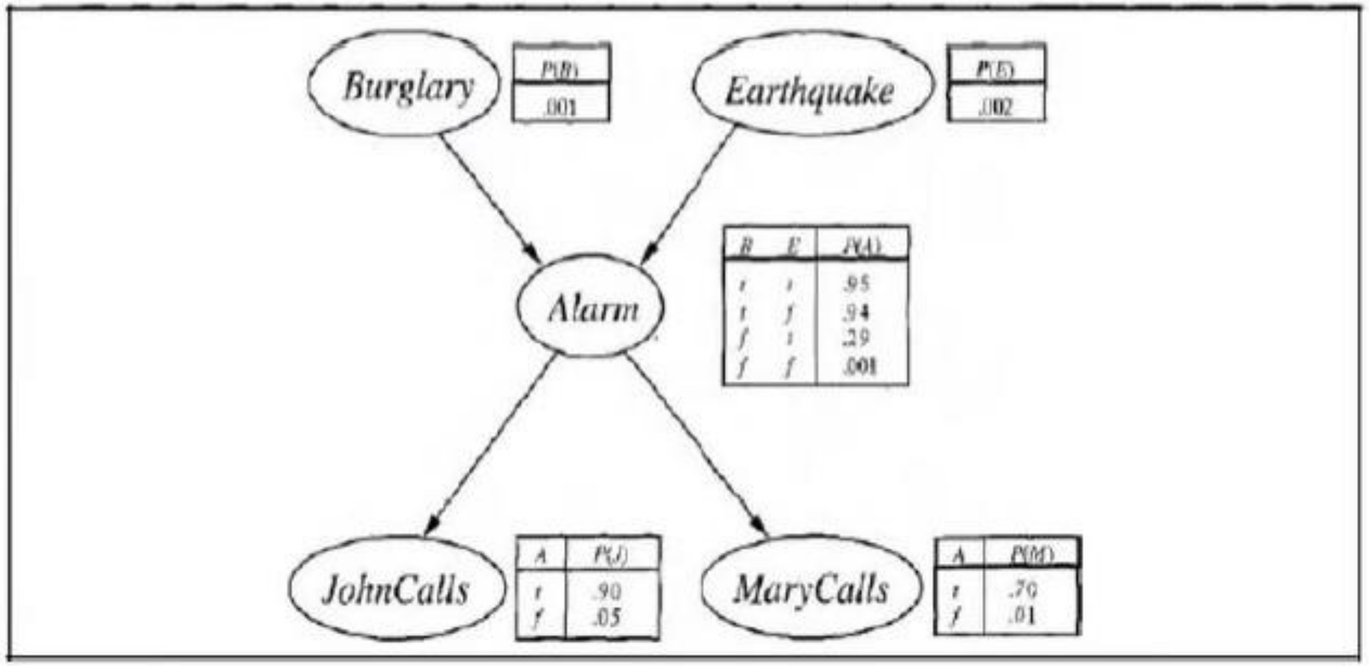
Figure 14.2 A typical Bayesian network, showing both the topology and the conditional probability tables (CPTs). In the CPTs, the letters B, E, A, J, and *M* stand for *Burglary*, *Earthquake*, *Alarm*, *JohnCalls*, and *MaryCalls*, respectively.

Example - Burglar alarm

- You have a new burglar alarm installed at home.
- It also responds on occasion to minor earthquakes.
- You also have two neighbors, John and Mary, they promised to call you when they hear the alarm.
- John always calls when he hears the alarm, but sometimes confuses the telephone ringing with the alarm, and calls then, too.
- Mary likes rather loud music, and sometimes, she misses the alarm altogether.
- Given the evidence of who has or has not called, we would like to estimate the probability of a burglary.

Bayesian Belief Network

- The burglary and earthquakes directly affect the probability of the alarm's going off,
- But, John and Mary call depends only on the alarm.
- The network does not have nodes for Mary's currently listening to loud music or the telephone ringing and confusing John.



Example

- We can calculate the probability that the alarm has sounded, but neither a burglary nor an earthquake has occurred, and both John and Mary call.

$$P(j \wedge m \wedge a \wedge \neg b \wedge \neg e)$$
$$= P(j|a)P(m|a)P(a|\neg b \wedge \neg e)P(\neg b)P(\neg e)$$
$$= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998 = 0.00062.$$

Semantics of Bayesian Network

- An entry in joint distribution is the probability of **conjunction** of particular assignment to each variable, such as
- $P(X1=x_1 \wedge X2=x_2 \wedge \ldots \wedge ..Xn=x_n)$ is equal to
- $$P(x_1,x_2,\ldots,x_n) = \prod_{i=1}^{n} P(x_i \mid Parent(X_i))$$

Method for Constructing Bayesian Network

- Rewrite the joint distribution in terms of a conditional probability, using the **product rule**

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

- Then we repeat the process, reducing each conjunctive probability to a conditional probability and a smaller conjunction. We end up with one big product:

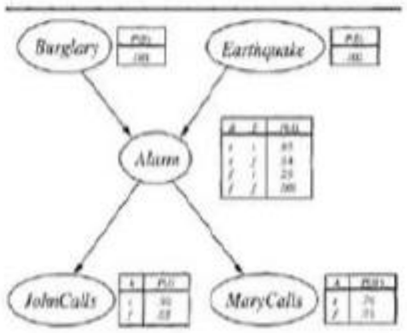
$$\begin{aligned} P(x_1, \ldots, x_n) &= P(x_n | x_{n-1}, \ldots, x_1) P(x_{n-1} | x_{n-2}, \ldots, x_1) \cdots P(x_2 | x_1) P(x_1) \\ &= \prod_{i=1}^{n} P(x_i | x_{i-1}, \ldots, x_1) \,. \end{aligned}$$

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

$$\mathbf{P}(MaryCalls | JohnCalls, Alarm, Earthquake, Burglary) = \mathbf{P}(MaryCalls | Alarm) \,.$$

Compactness and Node Ordering

- The compactness of Bayesian network is an example of general property of locally constructed systems. (also called as spare systems, inside some components there, and those are communicated)
- In a locally structured system, each subcomponent interacts directly with only a bounded number of other components, regardless of the total number of components.
- Therefore the correct order in which to add node is to add the 'root causes' first, then the variables they influenced and so on until we reach the leaves.



Bayesian Belief Network

- Suppose we decide to add the nodes in the order MaryCalls, JohnCalls, Alarm, Burglary, Earthquake.
- Adding Mary Calls: No parents.
- Adding JohnCalls: If Mary calls, that probably means the alarm has gone off, which of course would make it more likely that John calls. Therefore, JohnCalls needs Mary Calls as a parent
- Adding Alarm: Clearly, if both call, it is more likely that the alarm has gone off than if just one or neither call, so we need both MaryCalls and JohnCalls as parents.
- Adding Burglary: If we know the alarm state, then the call from John or Mary might give us information about our phone ringing or Mary's music, but not about burglary:
- **P(Burglary|Alarm, JohnCalls, MaryCalls) = P(BurglarylAlarm)**
- Hence we need just Alarm as parent.
- Adding *Earthquake:* if the alarm is on, it is more likely that there has been an earthquake. But if we know that there has been a burglary, then that explains the alarm, and the probability of an earthquake would be only slightly above normal. Hence, we need both *Alarm* and *Burglary* as parents.

Bayesian Belief Network

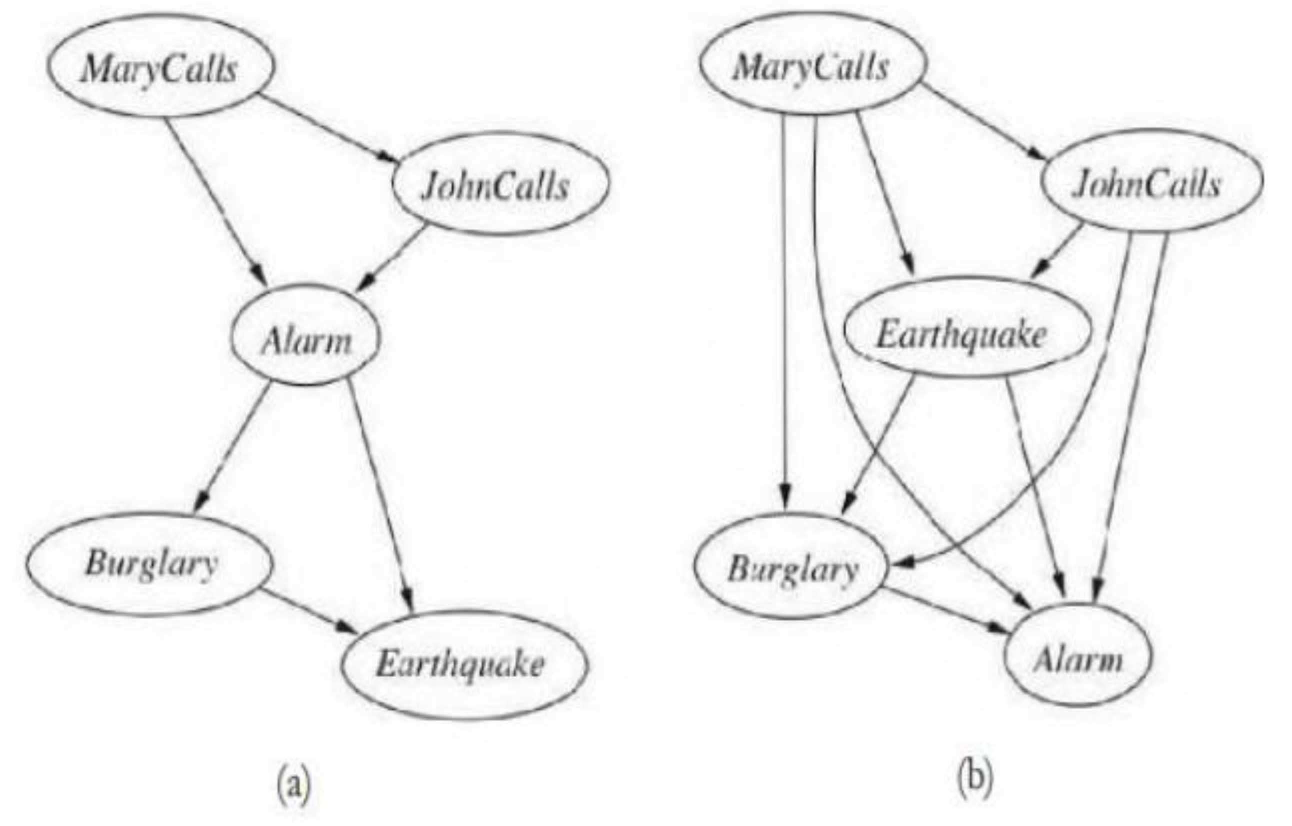
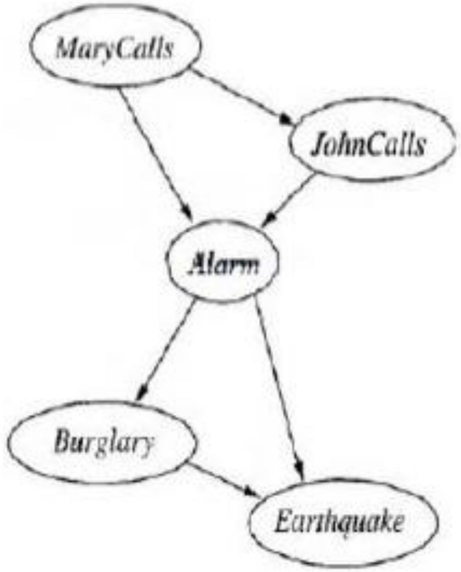


Figure 14.3 Network structure depends on order of introduction. In each network, we have introduced nodes in top-to-bottom order.

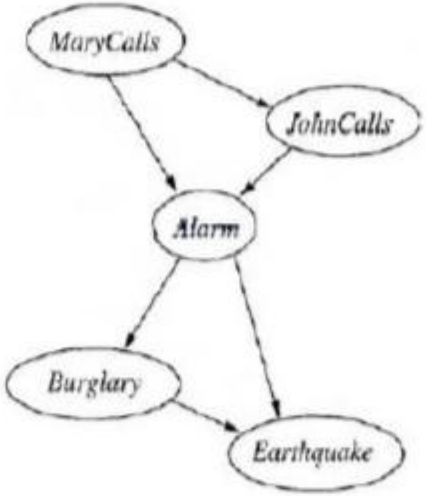
Conditional independence relations in Bayesian networks

- 1. A node is conditionally independent of its non-descendants, given its parents.
  - For example, *JohnCalls* is independent of .*Burglary* and *Earthquake*, given the value of *Alarm*.



Conditional independence relations in Bayesian networks...

- 2. A node is conditionally independent of all other nodes in the network, given its parents, children, and children's parents-that is, given its Markov blanket.
  - For example, *Burglary* is independent of *JohnCalls* and *MaryCalls*, given *Alarm* and *Earthquake*.



Bayesian Belief Network

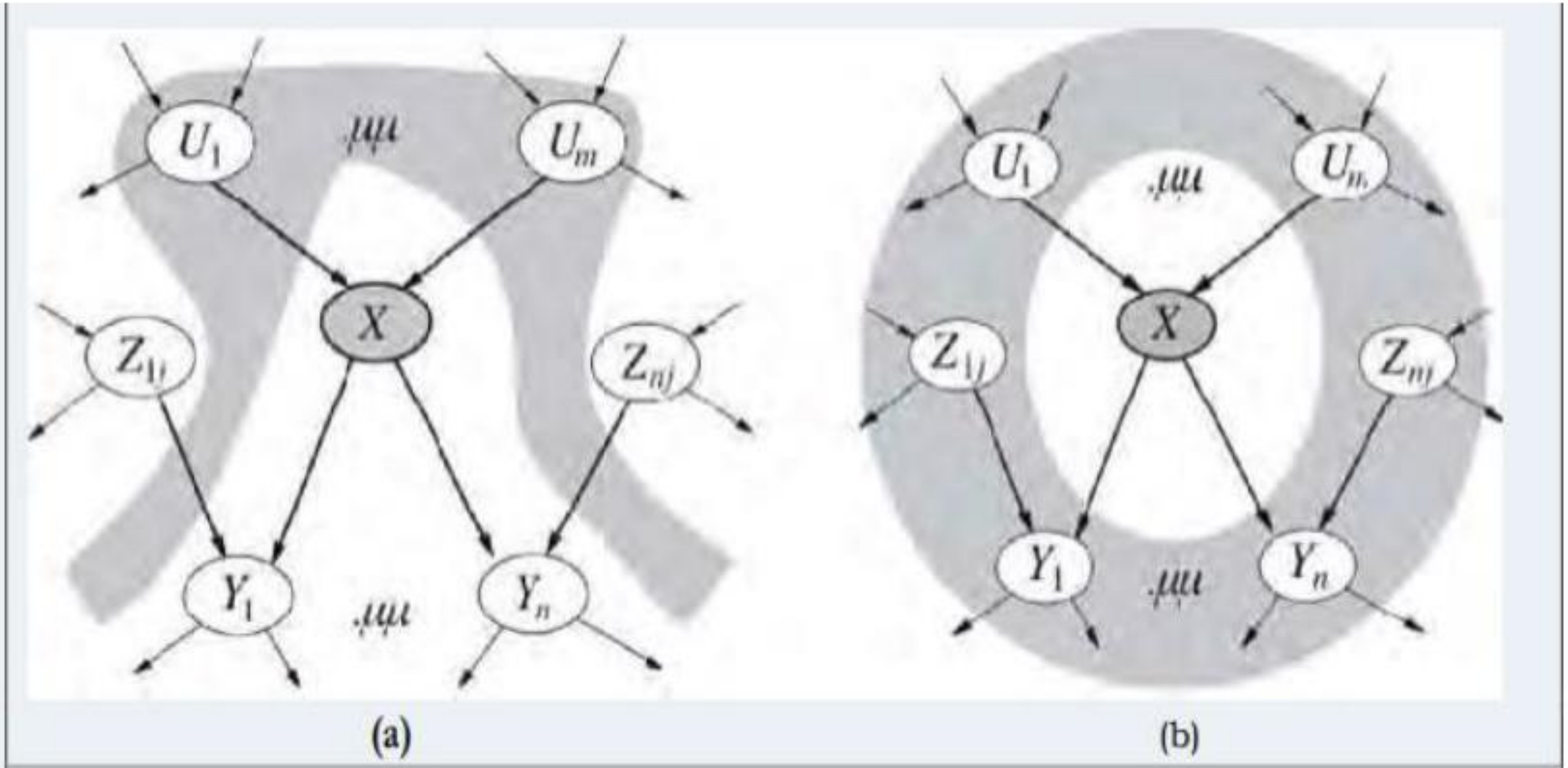


Figure 14.4 (a) A node X is conditionally independent of its non-descendants (e.g., the Z_{ij}s) given its parents (the U_is shown in the gray area). (b) A node X is conditionally independent of all other nodes in the network given its Markov blanket (the gray area).

Inferences in Bayesian Network - Purpose

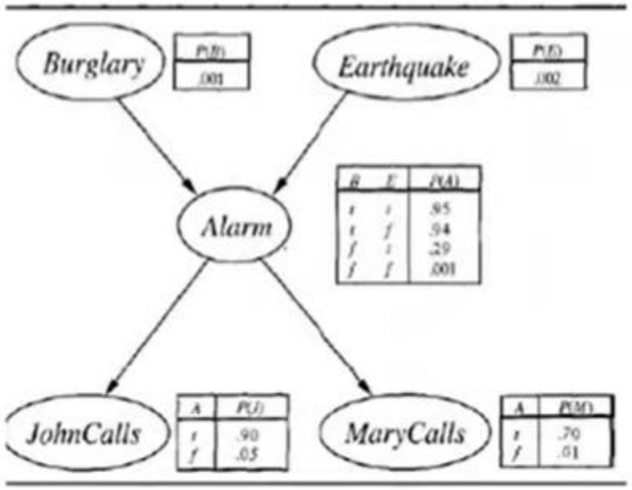
- Probabilistic Inference System is to compute Posterior Probability Distribution for a set of query variables, given some observed events.
- That is, some assignment of values to a set of evidence variables.

Inference in Bayesian Networks - Notations

- X – denotes the query variable
- E - set of evidence variables {E1, …. , Em}
- e – particular observed event
- Y – non-evidence, non-query variables, Y1, …, Yn. (Called the hidden variables)
- The complete set of variables - X = {X} U E U Y
- A typical query asks for the Posterior Probability Distribution **P(X|e)**

Approximate Inference Belief Networks

- In the burglary network, we might observe the event in which
  - *JohnCalls = true* and *MaryCalls = true.*
- We could then ask for, say, the probability that a burglary has occurred:
  - *P(Burglary | JohnCalls = true, MaryCalls = true) = (0.284,0.716)*



Types of Inferences

- Inference by Enumeration.
  - (inference by listing or recording all variables)
- Inference by Variable Elimination.
  - (inference by variable removal)

Inference by Enumeration

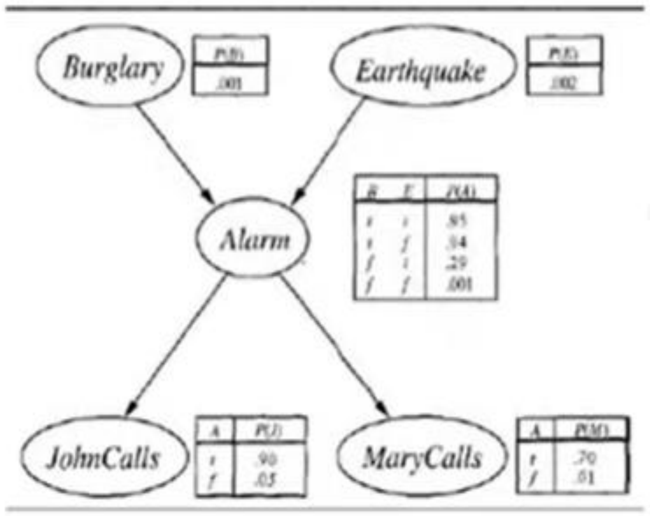
- Any conditional probability can be computed by summing terms from the full joint distribution.
- More specifically, a query P(X|e) can be answered using equation:

$$\mathbf{P}(X|\mathbf{e}) = \alpha\,\mathbf{P}(X,\mathbf{e}) = \alpha \sum_{\mathbf{Y}} \mathbf{P}(X,\mathbf{e},\mathbf{y})$$

- Where α is normalized constant
- X – Query Variable
- e – event
- y – number of terms

Inference by Enumeration...

- Consider P(Burglary | JohnCalls = true, MaryCalls = true)
- Burglary – query variable (X)
- JohnCalls – Evidence variable 1 (E1)
- MaryCalls – Evidence Variable 2 (E2)
- The hidden variables of this query are earthquake and alarm



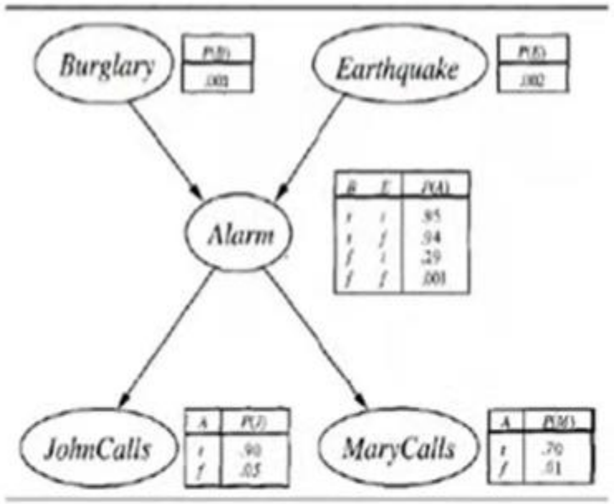
Inference by Enumeration...

- Using initial letter for the variables to shorten the expression we have

$$\mathbf{P}(B|j,m) = \alpha\,\mathbf{P}(B,j,m) = \alpha \sum_{e}\sum_{a}\mathbf{P}(B,e,a,j,m)$$

- The semantic of Bayesian network give us an expression, in terms of CPT entries, for simplicity we do this just for Burglary = true

$$P(b|j,m) = \alpha \sum_{e}\sum_{a} P(b)P(e)P(a|b,e)P(j|a)P(m|a)$$



Approximate Inference Belief Networks

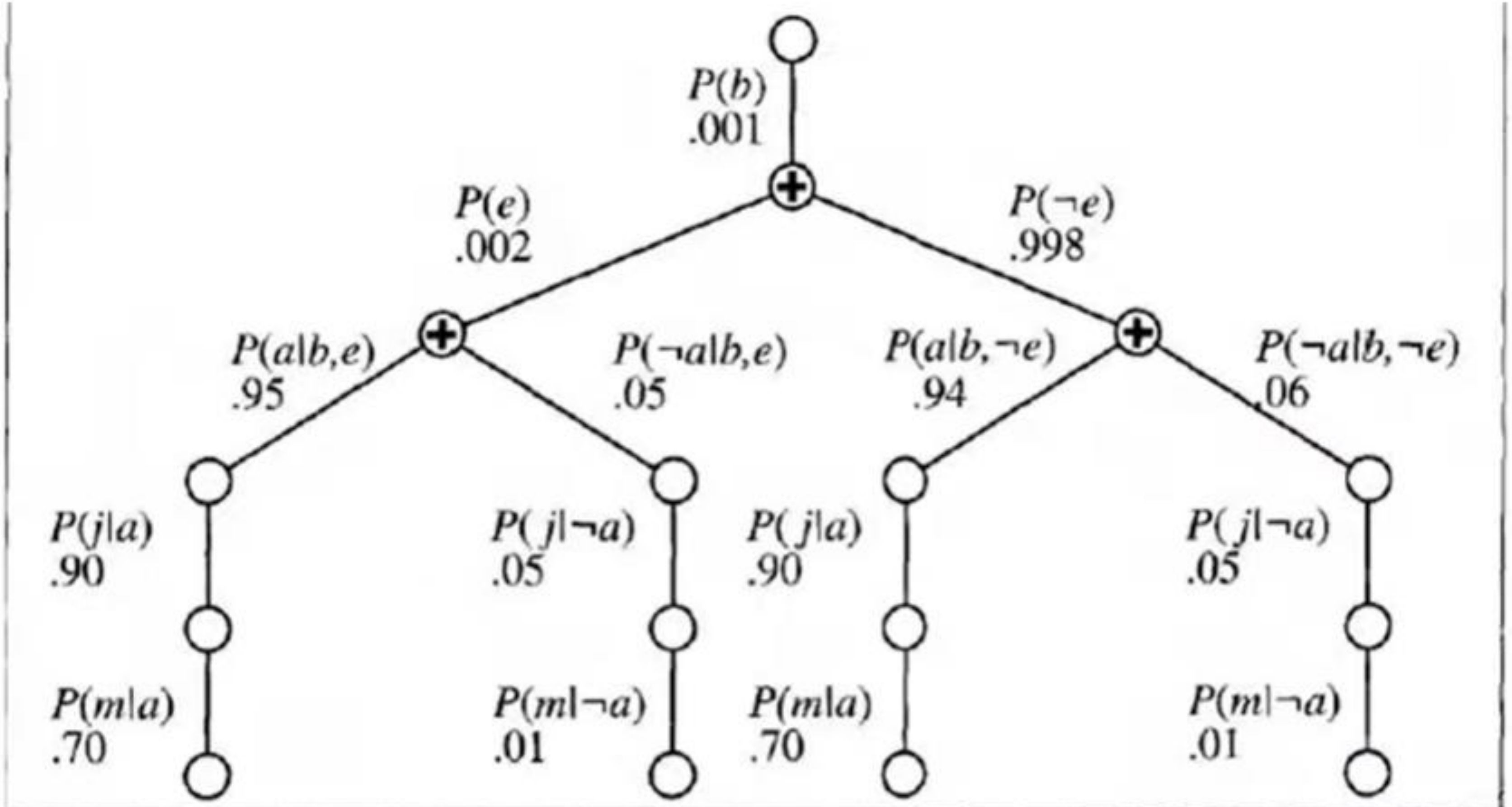


Figure 14.8 The structure of the expression shown in Equation (14.3). The evaluation proceeds top-down, multiplying values along each path and summing at the "+" nodes. Notice the repetition of the paths for *j* and m.

Inference by Variable Elimination

- The enumeration algorithm can be improved substantially by elimination repeated calculations.
- The idea is simple: do the calculation once and solve the result for later use. This is a form of dynamic programming.

Inference by Variable Elimination ...

- Variable elimination works by evaluating expressions,
- previous equation (derived in inference by enumeration)

$$P(b|j,m) = \alpha \sum_{e} \sum_{a} P(b)P(e)P(a|b,e)P(j|a)P(m|a)$$

- From this the repeated variables are separated

$$P(b|j,m) = \alpha P(b) \sum_{e} P(e) \sum_{a} P(a|b,e)P(j|a)P(m|a)$$

Inference by Variable Elimination ...

- Intermediate results are stored, and summations of each variable are done, for only those portion of the expression, that depends on the variable.
- Let us illustrate this process for the burglary network.
- We evaluate the expression

$$\mathbf{P}(B|j,m) = \alpha \underbrace{\mathbf{P}(B)}_{B} \sum_{e} \underbrace{P(e)}_{E} \sum_{a} \underbrace{\mathbf{P}(a|B,e)}_{A} \underbrace{P(j|a)}_{J} \underbrace{P(m|a)}_{M}$$

- We have annotated each part of the expression with the name of the associated variable; these parts are called factors

Inference by Variable Elimination ...

- For example, the factors f4(a) and f5(a) corresponding to P(j | a) and P(m|a) depending just on A because J and M are fixed by the query.
- They are therefore two element vectors.

$$\mathbf{f}_4(A) = \begin{pmatrix} P(j \mid a) \\ P(j \mid \neg a) \end{pmatrix} = \begin{pmatrix} 0.90 \\ 0.05 \end{pmatrix} \qquad \mathbf{f}_5(A) = \begin{pmatrix} P(m \mid a) \\ P(m \mid \neg a) \end{pmatrix} = \begin{pmatrix} 0.70 \\ 0.01 \end{pmatrix}$$

Inference by Variable Elimination – Example

- Given two factors f1 (A, B) and f2(B, C) with probability distributions shown below, the pointwise product f1 x f2 = f3(A, B, C) has $2^{1+1+1}=8$:

A	B	$\mathbf{f}_1(A,B)$	B	C	$\mathbf{f}_2(B,C)$	A	B	C	$\mathbf{f}_3(A,B,C)$
T	T	.3	T	T	.2	T	T	T	$.3 \times .2$
T	F	.7	T	F	.8	T	T	F	$.3 \times .8$
F	T	.9	F	T	.6	T	F	T	$.7 \times .6$
F	F	.1	F	F	.4	T	F	F	$.7 \times .4$
						F	T	T	$.9 \times .2$
						F	T	F	$.9 \times .8$
						F	F	T	$.1 \times .6$
						F	F	F	$.1 \times .4$

Note for Students

- **This power point presentation is for lecture, therefore it is suggested that also utilize the text books and lecture notes.**
- **Also Refer the solved and unsolved examples of Text and Reference Books.**