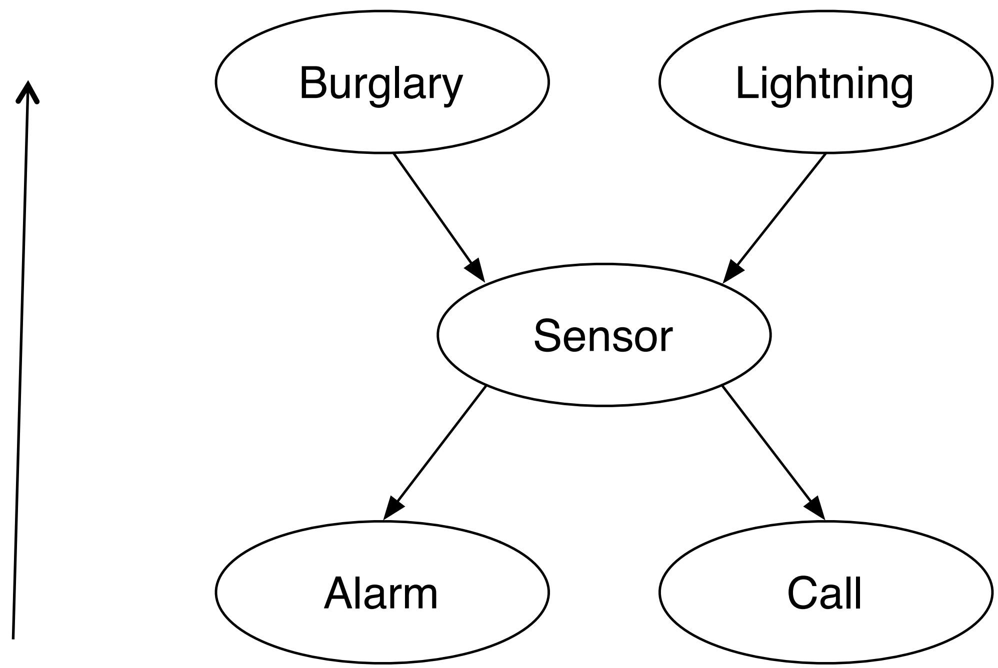
Bayesian Networks
Part 2 of 4
Defining probability equations

Compute P(alarm)

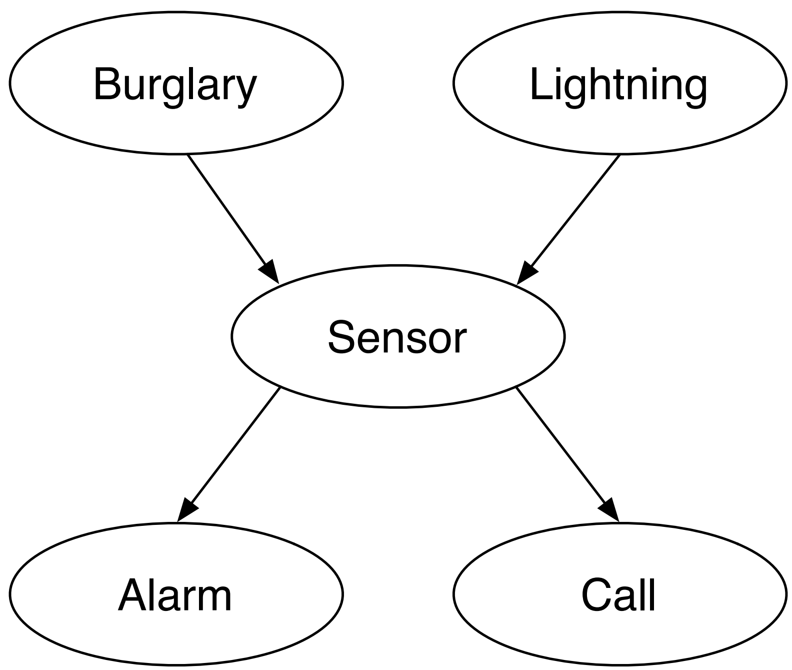
◊ The computation uses backward chained reasoning

» **Recall we have the probability tables for every event**



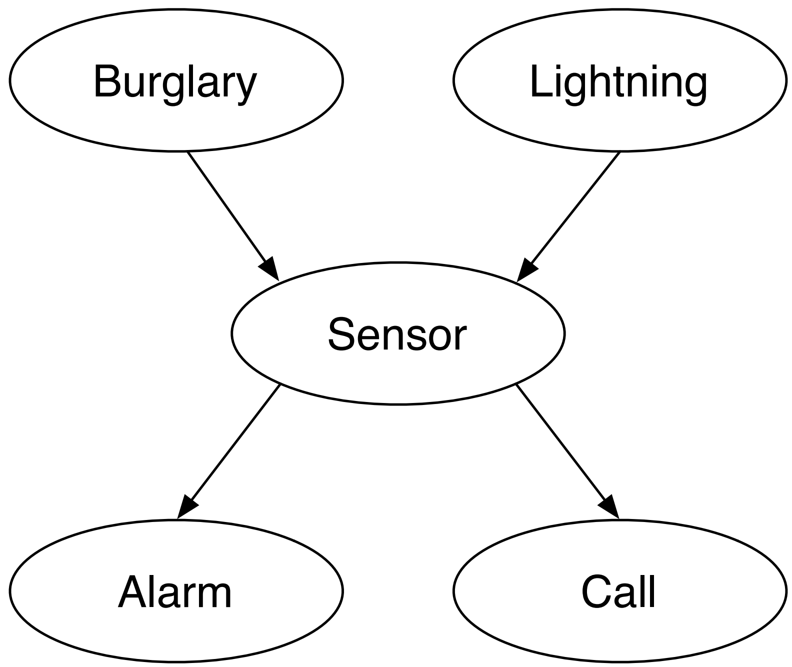
Compute P(alarm) – 2

- ◊ The computation uses backward chained reasoning
  - » **The probability computations (arithmetic) are in the forward direction**



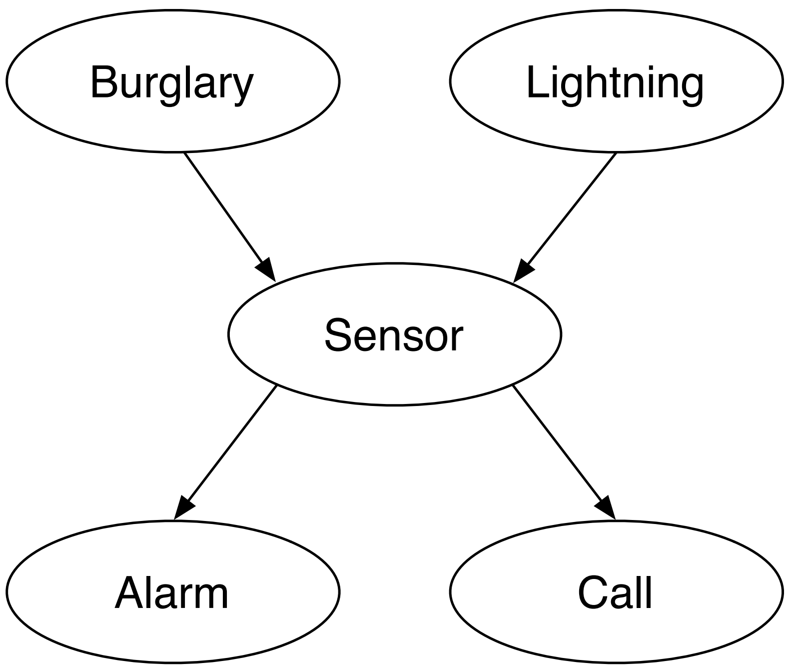
Compute P(alarm) – 3

- The computation uses backward chained reasoning
  - **The probability computations (arithmetic) are in the forward direction**
    - **From known values to unknown values.**



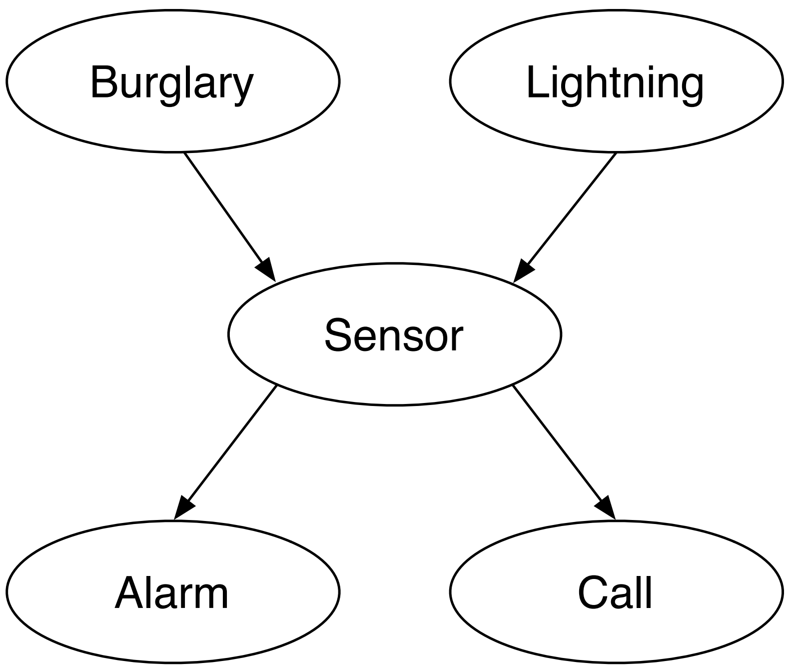
P(alarm) equation

- What is the probability equation for P(alarm)?



P(alarm) equation – 2

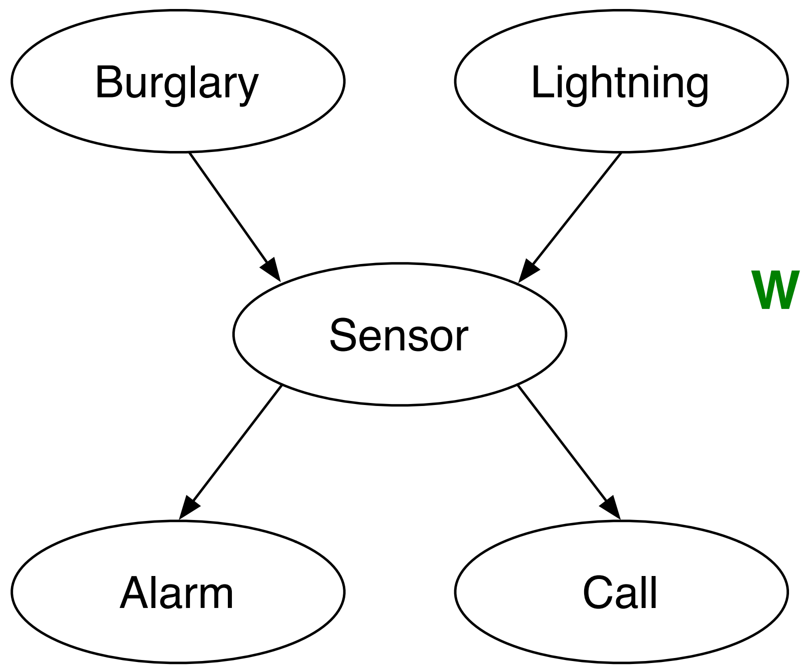
- The probability equation
  - » $P(alarm) = P(alarm \mid sensor) * P(sensor)$
    $+ P(alarm \mid \sim sensor) * P(\sim sensor)$



P(alarm) equation – 3

◊ The probability equation

» **P(alarm) = P(alarm | sensor) * P(sensor) + P(alarm | ~ sensor) * P(~ sensor)**



What do we do?

P(alarm) evaluation

◊ The probability equation

» P(alarm) = P(alarm I sensor) * P(sensor)
+ P(alarm I ~ sensor) * P(~ sensor)

> Evaluate the right hand side

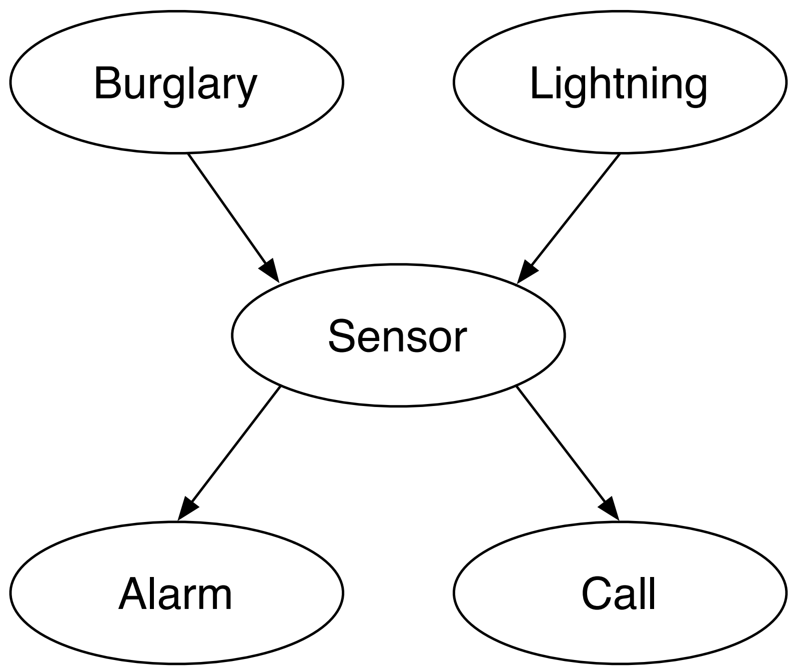
P(alarm) evaluation – 2

◊ The probability equation

» P(alarm) = P(alarm I sensor) * P(sensor)
+ P(alarm I ~ sensor) * P(~ sensor)

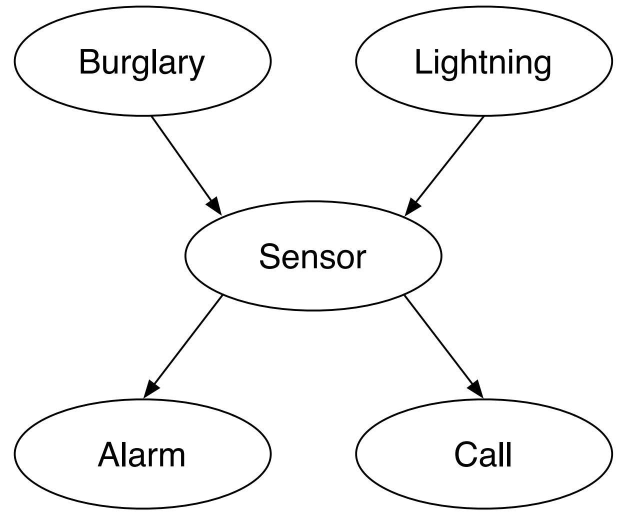
> Evaluate the right hand side

» So what do we know?
What are we given?



Burglar probability tables

» **So what do we know?**
What are we given?



◊ P(burglary) = 0.001
P(lightning) = 0.02

P(sensor | burglary, lightning) = 0.9
P(sensor | burglary, ~ lightning) = 0.9
P(sensor | ~ burglary, lightning) = 0.1
P(sensor | ~ burglary, ~ lightning) = 0.001

P(alarm | sensor) = 0.95
P(alarm | ~ sensor) = 0.001

P(call | sensor) = 0.9
P(call | ~ sensor) = 0.0

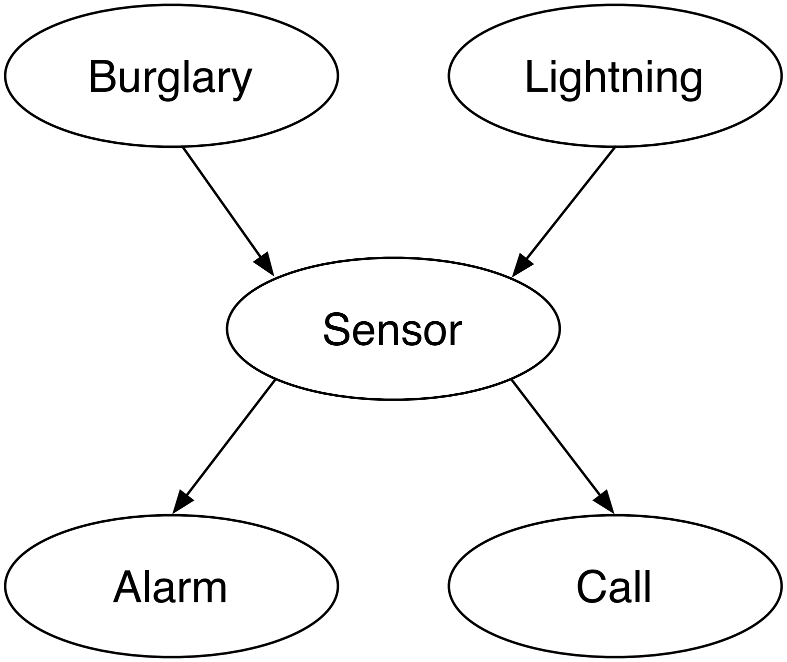
P(alarm) evaluation – 3

◊ The probability equation

» P(alarm) = P(alarm I sensor) * P(sensor)
    + P(alarm I ~ sensor) * P(~ sensor)

> Evaluate the right hand side

» Use what we are given



P(alarm) evaluation – 4

◊ The probability equation

» P(alarm) = P(alarm I sensor) * P(sensor)
+ P(alarm I ~ sensor) * P(~ sensor)

> Evaluate the right hand side

» P(alarm I sensor) = 0.95 given

P(alarm) evaluation – 5

◊ The probability equation

» P(alarm) = P(alarm I sensor) * P(sensor)
+ P(alarm I ~ sensor) * P(~ sensor)

> Evaluate the right hand side

» P(alarm I sensor) = 0.95 given

» P(alarm I ~ sensor) = 0.001 given

P(alarm) evaluation – 6

◊ The probability equation

» P(alarm) = P(alarm I sensor) * P(sensor)
+ P(alarm I ~ sensor) * P(~ sensor)

> Evaluate the right hand side

» P(alarm I sensor) = 0.95 given

» P(alarm I ~ sensor) = 0.001 given

» P(~ sensor) = 1 – P(sensor) probability rule

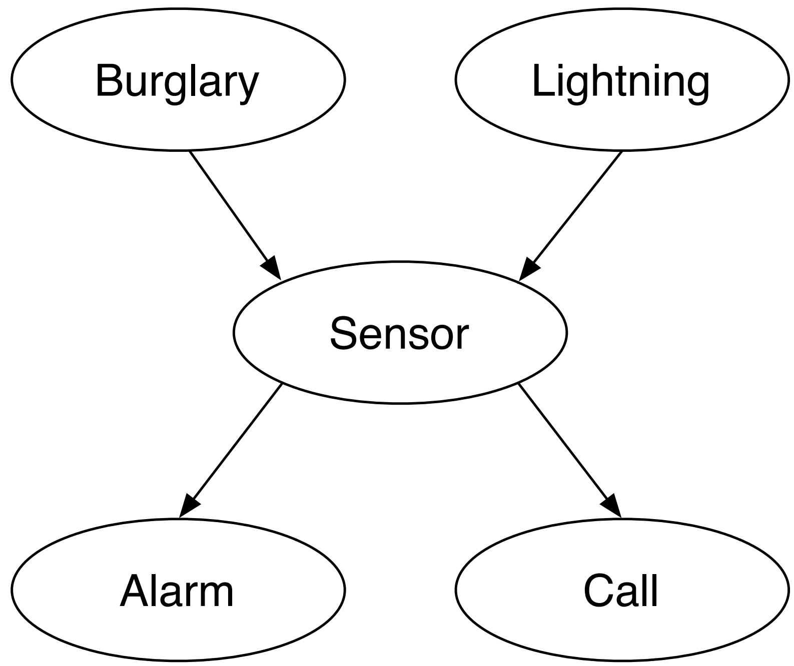
> What about P(sensor)?

P(alarm) evaluation – 7

- The probability equation
  - » P(alarm) = P(alarm | sensor) * P(sensor)
    + P(alarm | ~ sensor) * P(~ sensor)
    - > Evaluate the right hand side
  - » P(alarm | sensor) = 0.95 given
  - » P(alarm | ~ sensor) = 0.001 given
  - » P(~ sensor) = 1 – P(sensor) probability rule
    - > What about P(sensor)?
      - – **Need to compute it**

P(sensor) equation

» **What is the probability equation for P(sensor) ?**



P(sensor) equation – 2

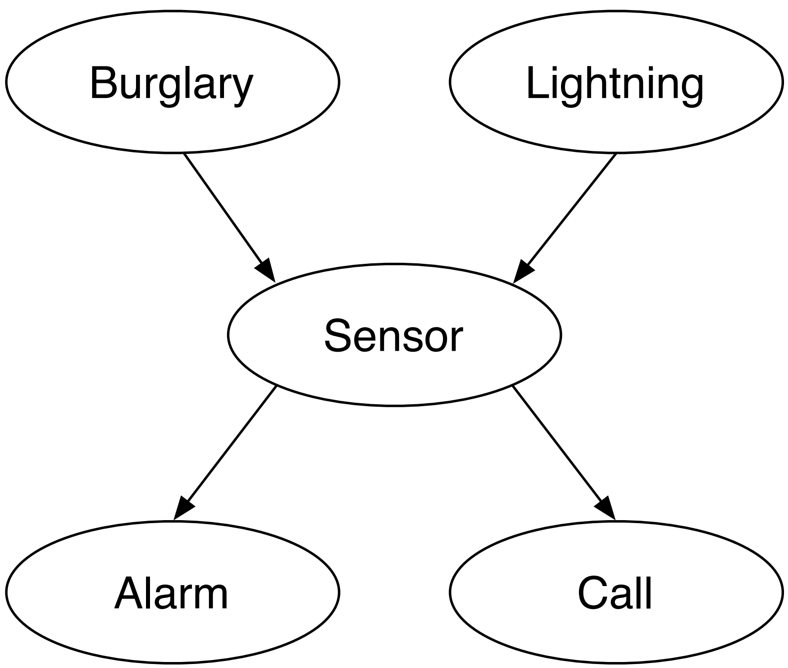
◊ The probability equation

» P(sensor) = P(sensor | burglary ^ lightning)
* P(burglary ^ lightning)

+ P(sensor | burglary ^ ~ lightning)
* P(burglary ^ ~ lightning)

+ P(sensor | ~ burglary ^ lightning)
* P(~ burglary ^ lightning)

+ P(sensor | ~ burglary ^ ~ lightning)
* P(~ burglary ^ ~ lightning)



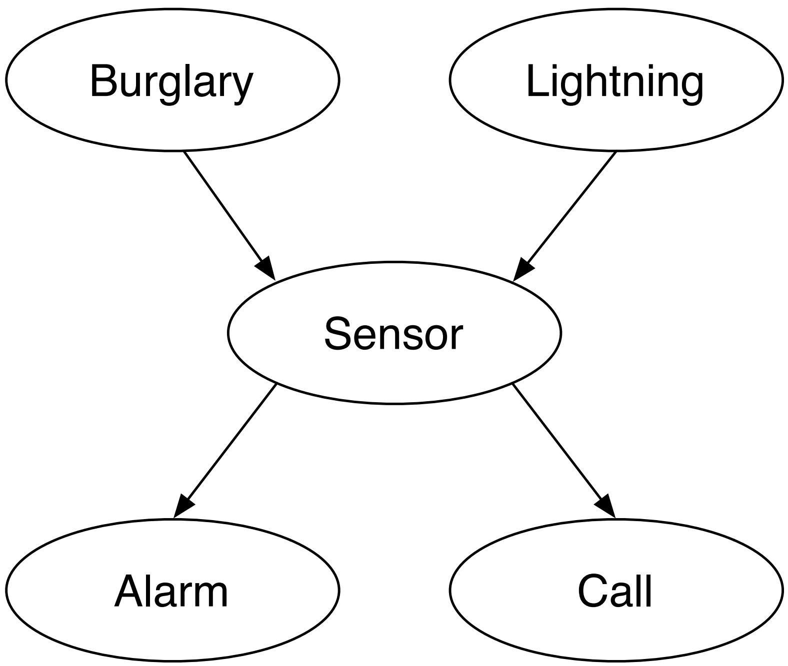
P(sensor) equation – 3

- The probability equation
  - » P(sensor) = P(sensor | burglary ^ lightning)
    * P(burglary ^ lightning)

    + P(sensor | burglary ^ ~ lightning)
    * P(burglary ^ ~ lightning)

    + P(sensor | ~ burglary ^ lightning)
    * P(~ burglary ^ lightning)

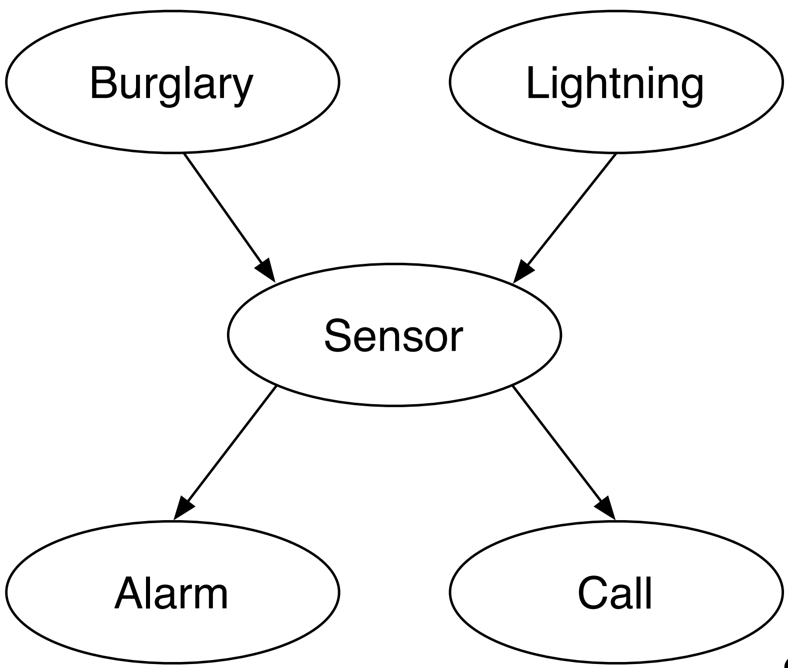
    + P(sensor | ~ burglary ^ ~ lightning)
    * P(~ burglary ^ ~ lightning)



What do we do?

P(sensor) equation – 4

- The probability equation
  - P(sensor) = P(sensor | burglary ^ lightning) * P(burglary ^ lightning)
    + P(sensor | burglary ^ ~ lightning) * P(burglary ^ ~ lightning)
    + P(sensor | ~ burglary ^ lightning) * P(~ burglary ^ lightning)
    + P(sensor | ~ burglary ^ ~ lightning) * P(~ burglary ^ ~ lightning)



What do we do?

– **Simplify the equation**

– **How?**

P(sensor) equation – 6

◊ Burglary and lightning are independent

> **So we get the following**

» **P(sensor) = P(sensor | burglary ^ lightning) * P(burglary) * P(lightning)**

+ P(sensor | burglary ^ ~ lightning) * P(burglary) * P(~ lightning)

+ P(sensor | ~ burglary ^ lightning) * P(~ burglary) * P(lightning)

+ P(sensor | ~ burglary ^ ~ lightning) * P(~ burglary) * P(~ lightning)

P(sensor) equation – 7

◊ Burglary and lightning are independent

> **So we get the following**

» **P(sensor) = P(sensor I burglary ^ lightning)**
*** P(burglary) * P(lightning)**

+ P(sensor I burglary ^ ~ lightning)
*** P(burglary) * P(~ lightning)**

+ P(sensor I ~ burglary ^ lightning)
*** P(~ burglary) * P(lightning)**

+ P(sensor I ~ burglary ^ ~ lightning)
*** P(~ burglary) * P(~ lightning)**

> **What do we do now?**

P(sensor) equation – 8

◊ Burglary and lightning are independent

> **So we get the following**

» **P(sensor) = P(sensor | burglary ^ lightning)**
*** P(burglary) * P(lightning)**

+ P(sensor | burglary ^ ~ lightning)
*** P(burglary) * P(~ lightning)**

+ P(sensor | ~ burglary ^ lightning)
*** P(~ burglary) * P(lightning)**

+ P(sensor | ~ burglary ^ ~ lightning)
*** P(~ burglary) * P(~ lightning)**

> **What do we do now?**

– **Substitute the values from the Burglary Probability Table**

P(sensor) evaluation

◊ Substitute the given probabilities

» **P(sensor) = 0.9 * 0.001 * 0.02** **(= 0.00001800)**
+ 0.9 * 0.001 * 0.98 **(= 0.00088200)**
+ 0.1 * 0.999 * 0.02 **(= 0.00199800)**
+ 0.001 * 0.999 * 0.98 **(= 0.00097902)**

= 0.00387702

P(sensor) evaluation – 2

◊ Substitute the given probabilities

» P(sensor) = 0.9 * 0.001 * 0.02 (= 0.00001800)
+ 0.9 * 0.001 * 0.98 (= 0.00088200)
+ 0.1 * 0.999 * 0.02 (= 0.00199800)
+ 0.001 * 0.999 * 0.98 (= 0.00097902)

= 0.00387702

> Why did we compute P(sensor)?

> What do we do now?

P(alarm) evaluation – 7

- ◊ The probability equation
  - » P(alarm) = P(alarm I sensor) * P(sensor)
    + P(alarm I ~ sensor) * P(~ sensor)
    - > Evaluate the right hand side
  - » P(alarm I sensor) = 0.95 given
  - » P(alarm I ~ sensor) = 0.001 given
  - » P(~ sensor) = 1 – P(sensor) probability rule
    - > What about P(sensor)?
  - » P(sensor) = 0.00387702

P(alarm) evaluation – 8

◊ The probability equation

» P(alarm) = P(alarm I sensor) * P(sensor)
+ P(alarm I ~ sensor) * P(~ sensor)

> Evaluate the right hand side

» P(alarm I sensor) = 0.95 given

» P(alarm I ~ sensor) = 0.001 given

» P(~ sensor) = 1 – P(sensor) probability rule

> What about P(sensor)?

» P(sensor) = 0.00387702

> What do we do now?

P(alarm) evaluation – 9

◊ The probability equation

» P(alarm) = P(alarm I sensor) * P(sensor)
+ P(alarm I ~ sensor) * P(~ sensor)

> Evaluate the right hand side

» P(alarm I sensor) = 0.95 given

» P(alarm I ~ sensor) = 0.001 given

» P(~ sensor) = 1 – P(sensor) probability rule

> What about P(sensor)?

» P(sensor) = 0.00387702

> What do we do now?

– Substitute the known values in the probability equation

P(alarm) evaluation – 10

◊ Substitute the given probabilities

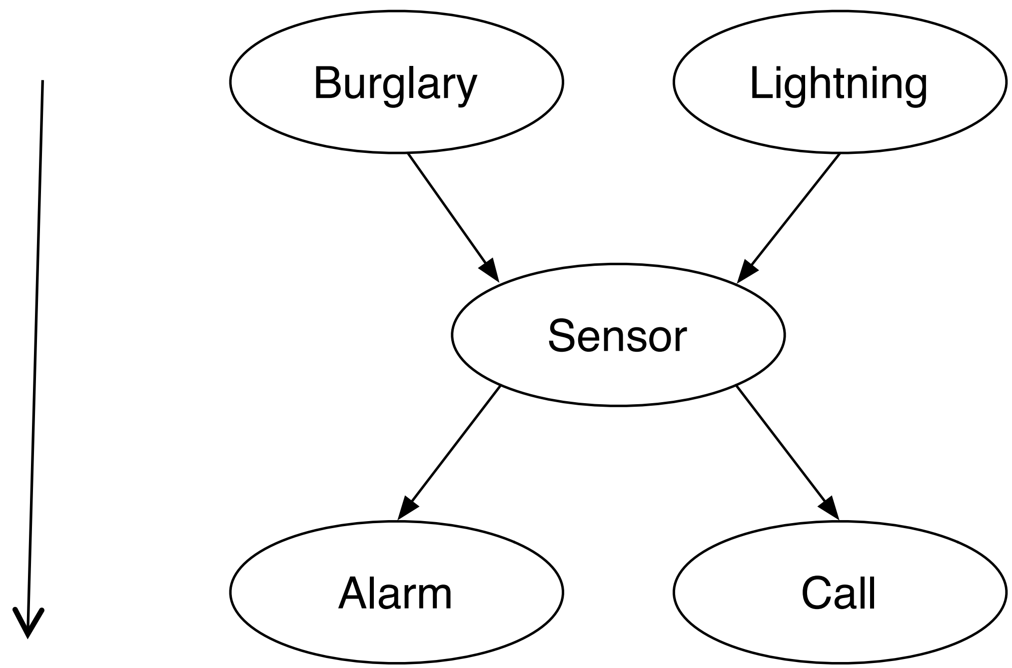
» P(alarm) = P(alarm I sensor) * P(sensor)
+ P(alarm I ~ sensor) * P(~ sensor)

» = 0.95 * 0.00387702 + 0.001 * 0.99612298

» = 0.00467929

P(burglary l alarm)

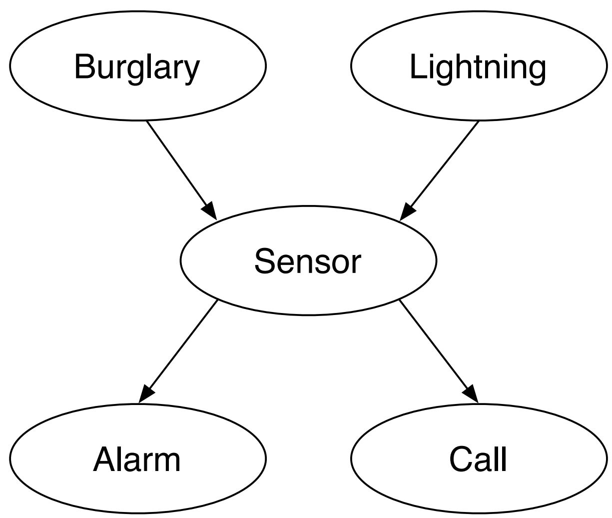
◊ The computation uses forward chained reasoning



P(burglary I alarm) – 2

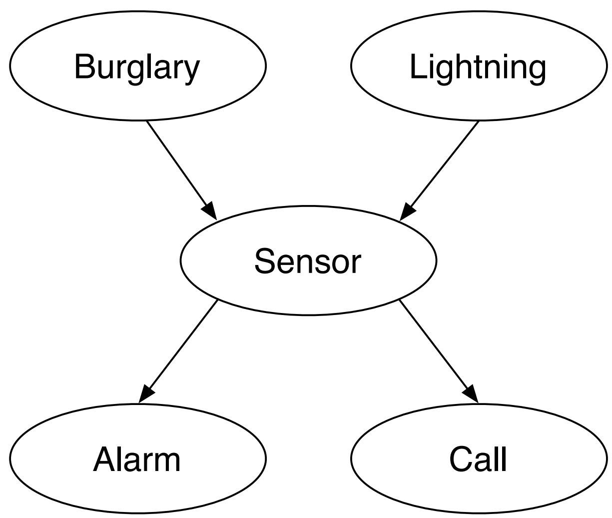
◊ The computation uses forward chained reasoning

» **The probability computations (arithmetic) can only be done in the forward direction**



P(burglary | alarm) – 3

- The computation uses forward chained reasoning
  - **The probability computations (arithmetic) can only be done in the forward direction**
    - **From known values to unknown values.**



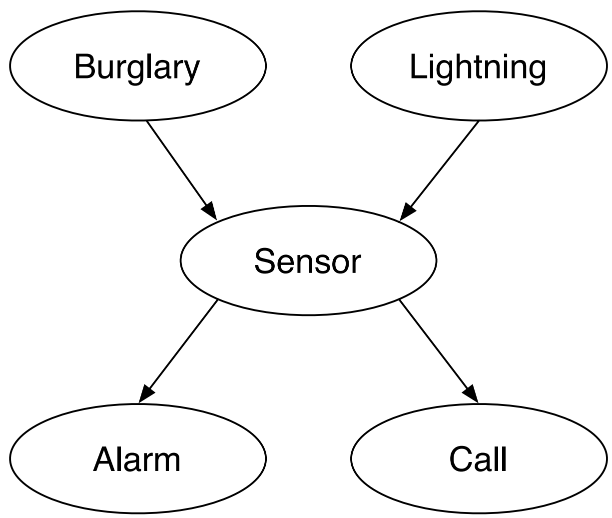
P(burglary I alarm) – 4

◊ The computation uses forward chained reasoning

» **But the probability computations (arithmetic) can only be done in the forward direction**

> **From known values to unknown values.**

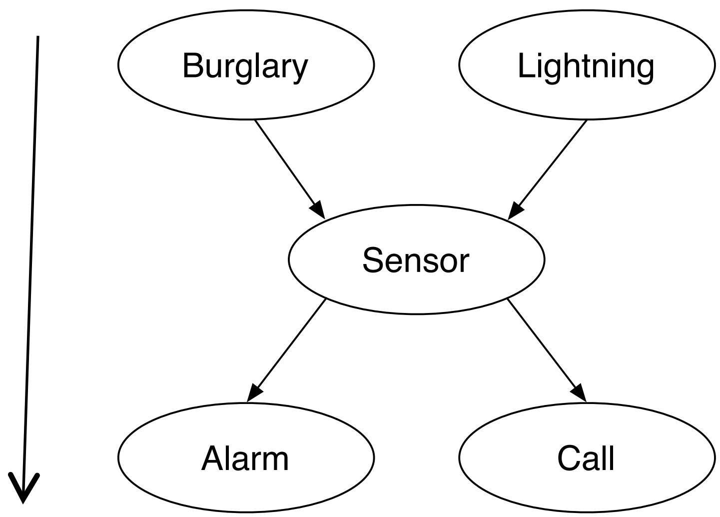
– **So what do we do?**



Bayes' formula

◊ Use Bayes' formula for forward chained reasoning

$$P(Hypothesis \mid Evidence) = P(Hypothesis) \frac{P(Evidence \mid Hypothesis)}{P(Evidence)}$$



Hypothesis = Burglary

Evidence = Alarm

Bayes' formula corollary

◊ Use Bayes' formula for forward chained reasoning

$$P(Hypothesis \mid Evidence) = P(Hypothesis) \frac{P(Evidence \mid Hypothesis)}{P(Evidence)}$$

◊ The following is a corollary

» **H = hypothesis   E = evidence   B = background knowledge**

$$P(H \mid E \wedge B) = P(H \mid B) \frac{P(E \mid H \wedge B)}{P(E \mid B)}$$

P(burglary | alarm) equation

» **What is the probability equation?**

P(burglary I alarm) equation – 3

» **What is the probability equation?**

◊ Use Bayes' formula

» **P(burglary I alarm) = P(burglary) * P(alarm I burglary) / P(alarm)**

$$P(\text{burglary} \mid \text{alarm}) = P(\text{burglary}) * \frac{P(\text{alarm} \mid \text{burglary})}{P(\text{alarm})}$$

> **Now what do we do?**

P(burglary | alarm) equation – 4

» **What is the probability equation?**

◊ Use Bayes' formula

» **P(burglary | alarm) = P(burglary) * P(alarm | burglary) / P(alarm)**

$$P(\text{burglary} \mid \text{alarm}) = P(\text{burglary}) * \frac{P(\text{alarm} \mid \text{burglary})}{P(\text{alarm})}$$

> **Now what do we do?**

– **Evaluate the RHS**

P(burglary | alarm) evaluation

- $P(burglary \mid alarm) = P(burglary) * \frac{P(alarm \mid burglary)}{P(alarm)}$
  - What do we know?

P(burglary | alarm) evaluation – 2

» $P(\text{burglary} \mid \text{alarm}) = P(\text{burglary}) * \dfrac{P(\text{alarm} \mid \text{burglary})}{P(\text{alarm})}$

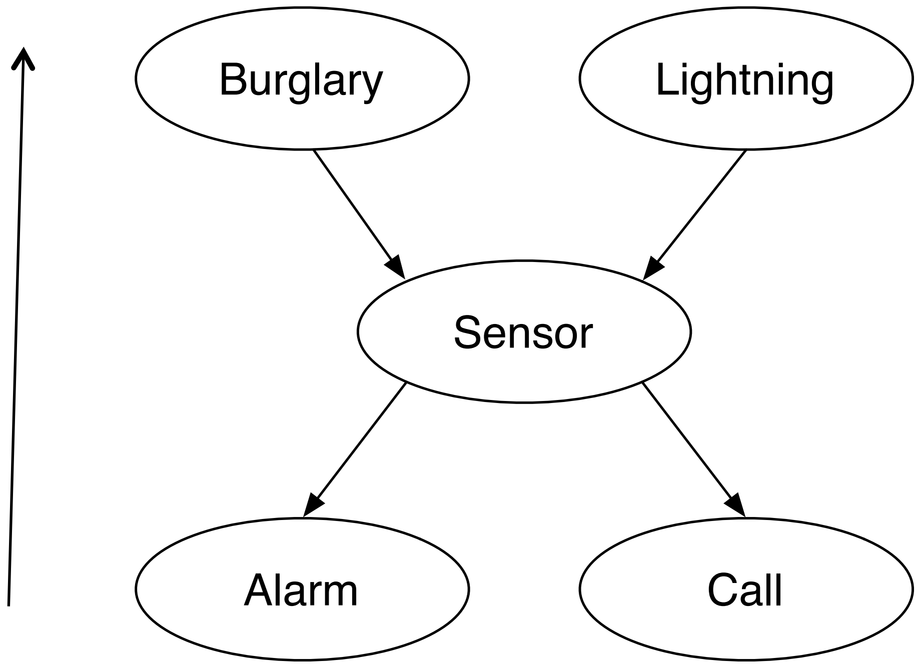
» P(burglary) = 0.001 given

» P(alarm) = 0.00467929 from an earlier computation

» P(alarm | burglary) = ??? need to compute

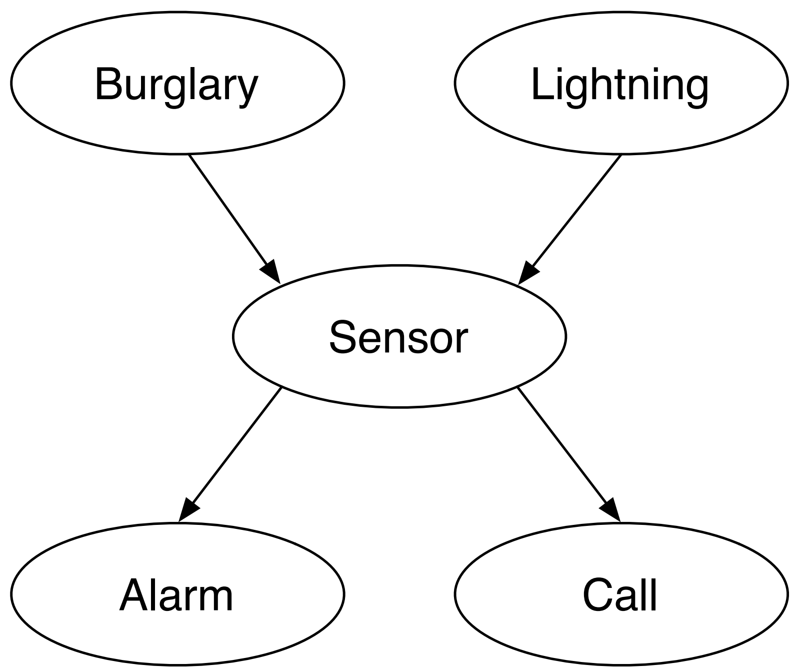
P(alarm | burglary) equation

- ◊ As before we use backward chained reasoning



P(alarm | burglary) equation – 2

- As before we use backward chained reasoning
  - **What is the probability equation?**

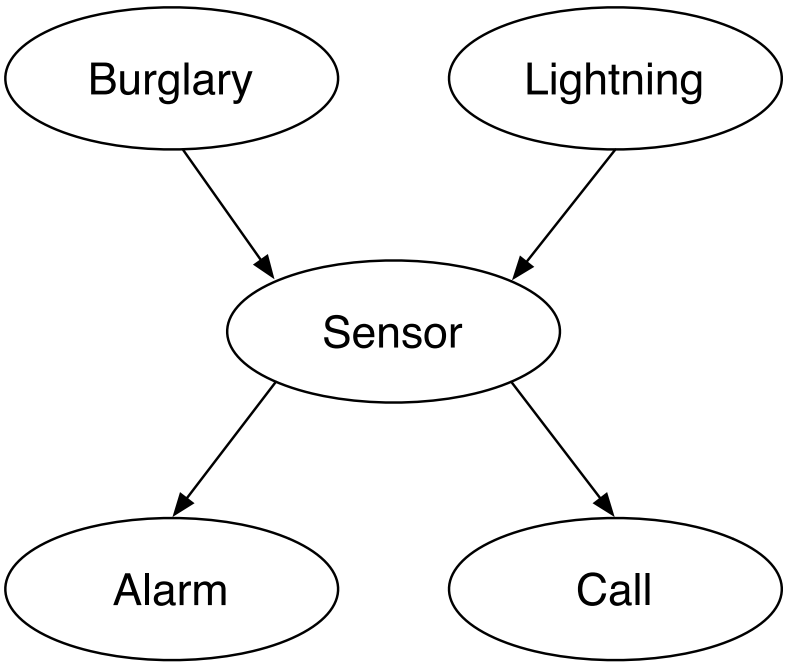


P(alarm | burglary) equation – 3

◊ As before we use backward chained reasoning

» **What is the probability equation?**

» **P(alarm | burglary) = P(alarm | sensor)**
*** P(sensor | burglary)**
+
P(alarm | ~ sensor)
*** P(~ sensor | burglary)**



P(alarm | burglary) evaluation

» P(alarm | burglary) = P(alarm | sensor)
  * P(sensor | burglary)
  +
  P(alarm | ~ sensor)
  * P(~ sensor | burglary)

  > What do we do?

P(alarm | burglary) evaluation – 2

» P(alarm | burglary) = P(alarm | sensor)
 * P(sensor | burglary)
 +
 P(alarm | ~ sensor)
 * P(~ sensor | burglary)

> What do we do?

– **Evaluate RHS by substituting known values**

P(alarm | burglary) evaluation – 3

» P(alarm | burglary) = P(alarm | sensor)
  * P(sensor | burglary)
  +
  P(alarm | ~ sensor)
  * P(~ sensor | burglary)

» P(alarm | sensor) = 0.95   given

P(alarm | burglary) evaluation – 4

» P(alarm | burglary) = P(alarm | sensor)
  * P(sensor | burglary)
  +
  P(alarm | ~ sensor)
  * P(~ sensor | burglary)

» P(alarm | sensor) = 0.95 given

» P(alarm | ~ sensor) = 0.001 given

P(alarm | burglary) evaluation – 5

- P(alarm | burglary) = P(alarm | sensor)
  * P(sensor | burglary)

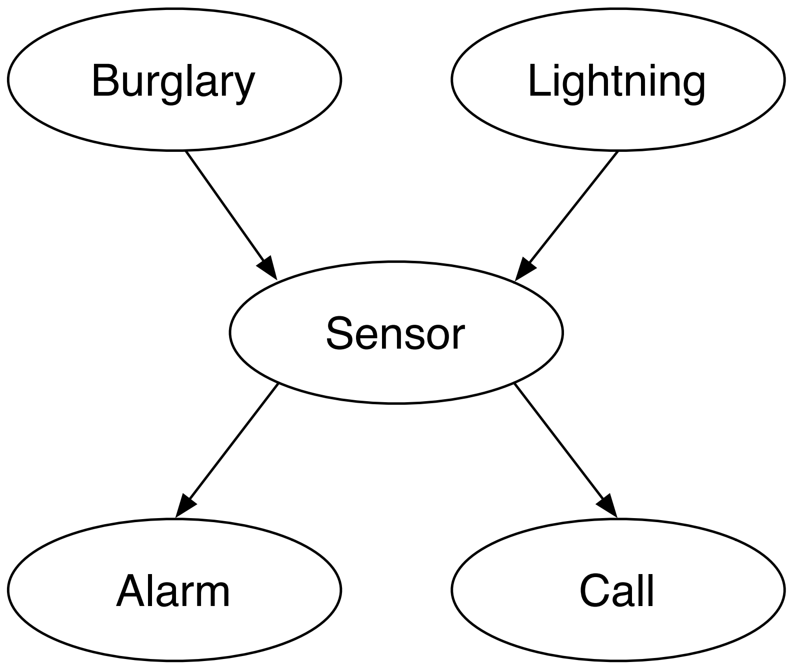
  +

  P(alarm | ~ sensor)
  * P(~ sensor | burglary)

- P(alarm | sensor) = 0.95 given
- P(alarm | ~ sensor) = 0.001 given
- P(sensor | burglary) = ??? need to compute

P(sensor | burglary) equation

» **What is the probability equation?**



P(sensor I burglary) equation – 2

◊ From an earlier slide we have

» P(sensor) = P(sensor I burglary ^ lightning)
  * P(burglary) * P(lightning)

  + P(sensor I burglary ^ ~ lightning)
  * P(burglary) * P(~ lightning)

  + P(sensor I ~ burglary ^ lightning)
  * P(~ burglary) * P(lightning)

  + P(sensor I ~ burglary ^ ~ lightning)
  * P(~ burglary) * P(~ lightning)

> What do we do now?

P(sensor | burglary) equation – 3

◊ From an earlier slide we have

» P(sensor) = P(sensor | burglary ^ lightning)
  * P(burglary) * P(lightning)

  + P(sensor | burglary ^ ~ lightning)
  * P(burglary) * P(~ lightning)

  + P(sensor | ~ burglary ^ lightning)
  * P(~ burglary) * P(lightning)

  + P(sensor | ~ burglary ^ ~ lightning)
  * P(~ burglary) * P(~ lightning)

> What do we do now?

– **Simplify the equation based on the evidence**

P(sensor | burglary) equation – 4

◊ Substitute P(burglary) = 1 and P(~ burglary) = 0

» **P(sensor) = P(sensor | burglary ^ lightning) * P(lightning)**

+ P(sensor | burglary ^ ~ lightning) * P(~ lightning)

> **What do we do now?**

P(sensor | burglary) equation – 5

◊ Substitute P(burglary) = 1 and P(~ burglary) = 0

» **P(sensor) = P(sensor | burglary ^ lightning)**
*** P(lightning)**

+ P(sensor | burglary ^ ~ lightning)
*** P(~ lightning)**

◊ Substitute the known values

» **P(sensor) = 0.9 * 0.02 + 0.9 * 0.98**
= 0.9

P(sensor | burglary) equation – 6

- ◊ Substitute P(burglary) = 1 and P(~ burglary) = 0
  - » **P(sensor) = P(sensor | burglary ^ lightning)**
    *** P(lightning)**

    + P(sensor | burglary ^ ~ lightning)
    *** P(~ lightning)**
- ◊ Substitute the known values
  - » **P(sensor) = 0.9 * 0.02 + 0.9 * 0.98**
    = 0.9

    > **What do we do now?**

P(alarm | burglary) evaluation – 7

» P(alarm | burglary) = P(alarm | sensor)
  * P(sensor | burglary)
  +
  P(alarm | ~ sensor)
  * P(~ sensor | burglary)

» P(alarm | sensor) = 0.95 given

» P(alarm | ~ sensor) = 0.001 given

» P(sensor | burglary) = 0.9 from computation

P(alarm | burglary) evaluation – 7

» P(alarm | burglary) = P(alarm | sensor)
* P(sensor | burglary)
+
P(alarm | ~ sensor)
* P(~ sensor | burglary)

» P(alarm | sensor) = 0.95 given

» P(alarm | ~ sensor) = 0.001 given

» P(sensor | burglary) = 0.9 from computation

> What is missing?
> How do we compute it?

P(alarm | burglary) evaluation – 8

» P(alarm | burglary) = P(alarm | sensor)
* P(sensor | burglary)
+
P(alarm | ~ sensor)
* P(~ sensor | burglary)

» P(alarm | sensor) = 0.95 given

» P(alarm | ~ sensor) = 0.001 given

» P(sensor | burglary) = 0.9 from computation

» P(~ sensor | burglary) = 1 – P(sensor | burglary)
= 0.1

P(alarm | burglary) evaluation – 9

- P(alarm | burglary) = P(alarm | sensor)
  * P(sensor | burglary)
  +
  P(alarm | ~ sensor)
  * P(~ sensor | burglary)

- P(alarm | sensor) = 0.95 given
- P(alarm | ~ sensor) = 0.001 given
- P(sensor | burglary) = 0.9 from computation
- P(~ sensor | burglary) = 1 – P(sensor | burglary)
  = 0.1
  - What do we do now?

P(alarm | burglary) evaluation – 10

» P(alarm | burglary) = P(alarm | sensor)
    * P(sensor | burglary)
  +
    P(alarm | ~ sensor)
    * P(~ sensor | burglary)

» P(alarm | sensor) = 0.95 given
» P(alarm | ~ sensor) = 0.001 given
» P(sensor | burglary) = 0.9 from computation
» P(~ sensor | burglary) = 1 – P(sensor | burglary)
  = 0.1

◊ Substitute the known values

» P(alarm | burglary) = 0.8551

P(burglary I alarm) evaluation – 3

- We have the following from “evaluation – 2”
  - » **P(burglary I alarm) = P(burglary) * $\frac{\text{P(alarm I burglary)}}{\text{P(alarm)}}$**
  - » **P(burglary) = 0.001** **given**
  - » **P(alarm) = 0.00467929** **from an earlier computation**
  - » **P(alarm I burglary) = ???** **need to compute**
- Now we know
  - » **P(alarm I burglary) = 0.8551**
  - » **P(burglary I alarm) = 0.001 * (0.8551 / 0.00467929)**
    = 0.1827414

Probability reasoning rules

- (1) Probability of a conjunction

$$P(N1 \wedge N2 \mid C) = P(N1 \mid C) \times P(N2 \mid N1 \wedge C)$$

- (2) Probability of a certain event

$$P(N \mid N \wedge C1 \wedge \ldots) = 1$$

- (3) Probability of an impossible event

$$P(N \mid \sim N \wedge C1 \wedge \ldots) = 0$$

- (4) Probability of a negation

$$P(\sim N \mid C) = 1 - P(N \mid C)$$

Probability reasoning rules – 2

◊ (5) If the condition involves a descendant of N use the corollary (general form) of Bayes' theorem

$$P(N \mid D \wedge C) = P(N \mid C) \times P(D \mid N \wedge C) / P(D \mid C)$$

◊ (6) If the condition does not involve a descendent of N then

◊ (6a) If N has no parents

$$P(N \mid C) = P(N)$$ **Given in the model**

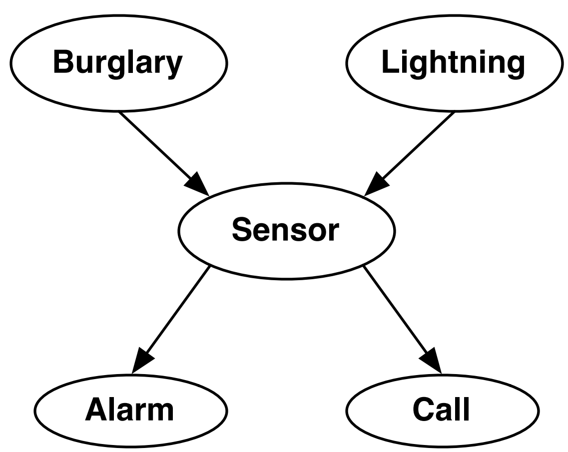
◊ (6b) If N has parents

$$P(N \mid C) = \sum_{S \in states(parents)} P(N \mid S) \times P(S \mid C)$$

P(call | alarm) equation

- The probability equation
  - P(call | alarm) = P(call | sensor) * P(sensor | alarm)
    + P(call | ~sensor) * P(~sensor | alarm)
    = 0.70841

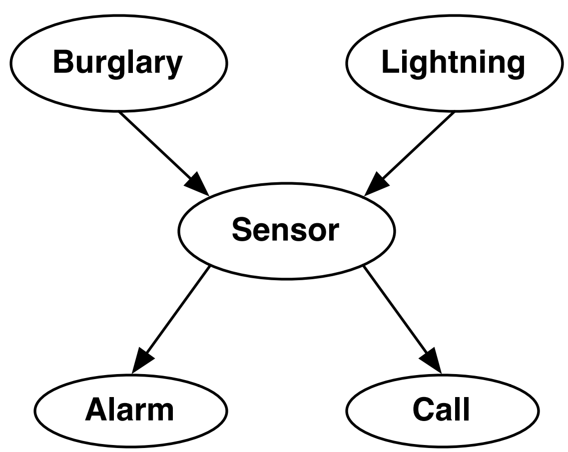
P(call | sensor) = 0.95 given
P(call | ~sensor) = 0.0 given
P(sensor | alarm) = 0.78712 see next slide
P(~sensor | alarm) = 1 – P(sensor | alarm) = 0.21288



P(sensor | alarm) equation

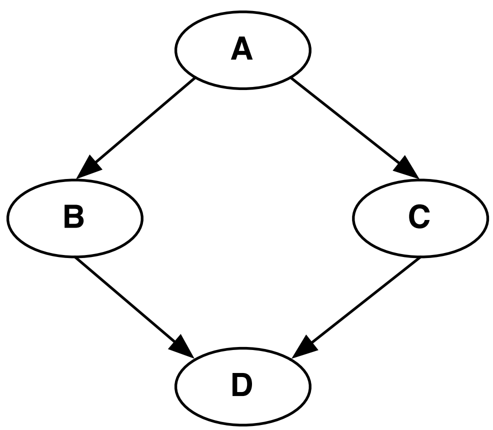
- The probability equation
  - P(sensor | alarm) = P(sensor) * P(alarm | sensor) / P(alarm)

    P(sensor) = 0.00387702 see previous slides
    P(alarm | sensor) = 0.95 given
    P(alarm) = 0.467929 see previous slides



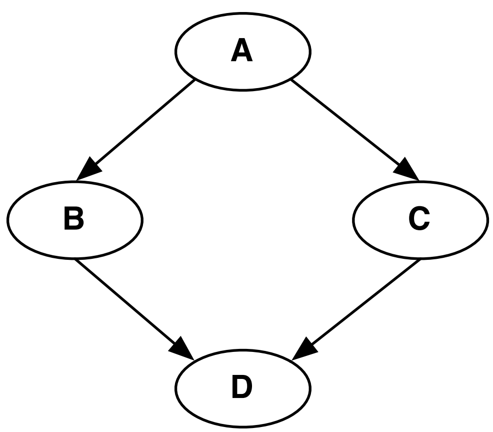
P(D) equation

- The probability equation continued
  - $P(D) = P(D \mid B, C) * P(B, C)$ $+ P(D \mid B, \sim C) * P(B, \sim C)$ $+ P(D \mid \sim B, C) * P(\sim B, C)$ $+ P(D \mid \sim B, \sim C) * P(\sim B, \sim C)$ — Given in the model
    - **Here B and C are dependent**
      - **Cannot simplify as we did with the P(sensor)**



P(D) equation – 2

- The probability equation
  - $P(B, C) = P(B) * P(C \mid B)$
  - $P(B, \sim C) = P(B) * P(\sim C \mid B)$
    $= P(B) * 1 - P(C \mid B)$
  - $P(\sim B, C) = P(\sim B) * P(C \mid \sim B)$
    $= (1 - P(B)) * P(C \mid \sim B)$
  - $P(\sim B, \sim C) = P(\sim B) * P(\sim C \mid \sim B)$
    $= (1 - P(B)) * (1 - P(C \mid \sim B))$
  - $P(B) = P(B \mid A) * P(A) + P(B \mid \sim A) * P(\sim A)$
    Given in the model



P(D) equation – 3

- The probability equation
  - P(C | B) = P(C | A) * P(A | B)
    + P(C | ~A) * P(~A | B) = P(C | ~A) * (1 – P(A | B))
  - P(C | ~B) = P(C | A) * P(A | ~B)
    + P(C | ~A) * P(~A | ~B) = P(C | ~A) * (1 – P(A | ~B))
  - P(A | B) = P(A) * P(B | A) / P(B)
  - P(A | ~B) = P(A) * P(~B | A) / P(~B)
    = P(A) * P(~B | A) / (1 – P(B))

Given in the model or computed previously