Quantifying Uncertainty & Probabilistic Reasoning

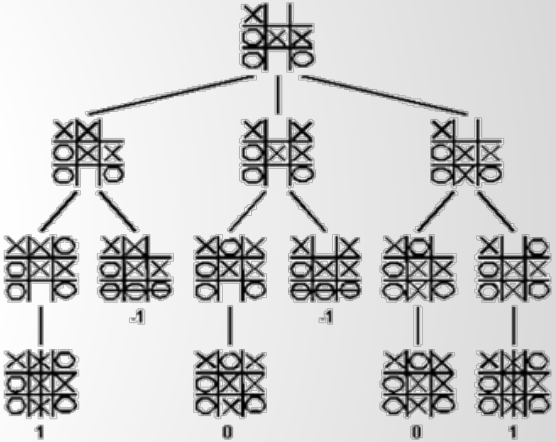
Abdulla AlKhenji
Khaled AlEmadi
Mohammed AlAnsari

Outline

- Previous Implementations
- What is Uncertainty?
- Acting Under Uncertainty
- Rational Decisions
- Basic Probability Notation
- Baye’s Rule
- Full Joint Distribution
- Probabilistic Reasoning

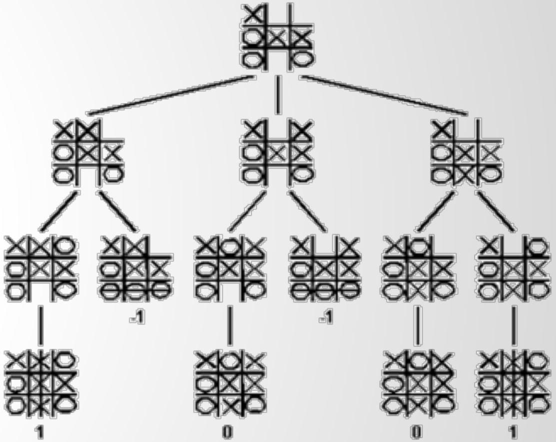
Previous Implementations

- Agents keep a track of a **belief state**, and generating a plan to act under every possible situation.



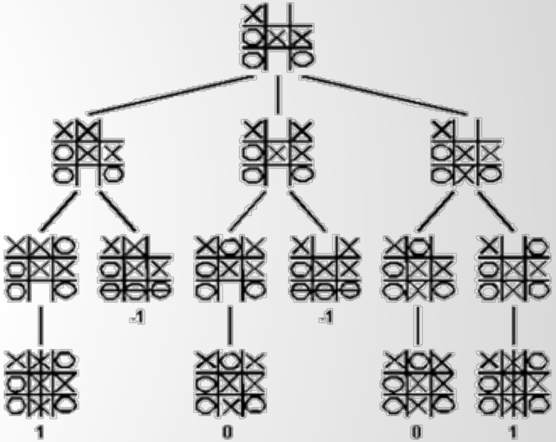
Previous Implementations

- Problems:
  - Logical Agent must consider every logically possible explanation
  - A correct contingency plan can grow arbitrarily large
  - Sometimes, no plan to achieve goal but agent has to act.



Previous Implementations

- Agents may need to handle **uncertainty**, whether due to partial observability, nondeterminism, or a combination of both.



What is Uncertainty?

- How early should I leave my house to reach my meeting?
- Should I take this road? is it going to be crowded?
- Should I play this card? What if he has a better card?
- I don’t know what to do… but I have to do something.

What is Uncertainty?

- Consider an automated Taxi-Driver, that has a plan A_{90} to leave the house 90 minutes before the departure and driving at a reasonable speed.
- Even if the airport is 5km away from the house, the agent *cannot* guarantee that it will reach on time.
- The plan's success cannot be inferred, leading to a **qualification problem**: no complete solution within logic.

Uncertain Reasoning

Example of Uncertain Reasoning:

- *Toothache* $\Rightarrow$ *Cavity*

Uncertain Reasoning

Example of Uncertain Reasoning:

- *Toothache* ⇒ *Cavity* (Fails)

Uncertain Reasoning

Example of Uncertain Reasoning:

- *Toothache* ⇒ *Cavity* (Fails)
- *Toothache* ⇒ *Cavity* V *GumProblem* V *Abscess*

Uncertain Reasoning

Example of Uncertain Reasoning:

- *Toothache* ⇒ *Cavity* (Fails)
- *Toothache* ⇒ *Cavity* V *GumProblem* V *Abscess* (Fails)

Uncertain Reasoning

Example of Uncertain Reasoning:

- *Toothache* ⇒ *Cavity* (Fails)
- *Toothache* ⇒ *Cavity V GumProblem V Abscess* (Fails)
- *Cavity* ⇒ *Toothache*

Uncertain Reasoning

Example of Uncertain Reasoning:

- *Toothache* ⇒ *Cavity* (Fails)
- *Toothache* ⇒ *Cavity V GumProblem V Abscess* (Fails)
- *Cavity* ⇒ *Toothache* (Fails)

Uncertain Reasoning

Example of Uncertain Reasoning:

- *Toothache* ⇒ *Cavity* (Fails)
- *Toothache* ⇒ *Cavity V GumProblem V Abscess* (Fails)
- *Cavity* ⇒ *Toothache* (Fails)

Reasons of Failure:

- Laziness: too many antecedents or consequents.
- Theoretical Ignorance: no complete theory for domain.
- Practical Ignorance: not all tests can or have been run.

Uncertain Reasoning

Example of Uncertain Reasoning:

- *Toothache* ⇒ *Cavity* (Fails)
- *Toothache* ⇒ *Cavity V GumProblem V Abscess* (Fails)
- *Cavity* ⇒ *Toothache* (Fails)

There is **no logical consequence** between Toothache in either direction. Same applies for ‘judgemental domains’ including: Business, Law, Design, Auto Repair etc…

Acting Under Uncertainty

The Agent's knowledge provides only a **Degree of Belief**.

The best way to deal with degrees of belief: **Probability Theory**.

Acting Under Uncertainty

The **ontological commitments** used previously are the same, but the **epistemological commitments** are different:

Acting Under Uncertainty

The **ontological commitments** used previously are the same, but the **epistemological commitments** are different:

- A **Logical Agent** believes a sentence is either *True* or *False*
- A **Probabilistic Agent** has a **degree of belief between 0** and **1**

(0 = False, 1 = True)

Acting Under Uncertainty

“Probability provides a way of summarizing the uncertainty that comes from our laziness and ignorance, thereby solving the qualification problem.”

Acting Under Uncertainty

<u>One confusing point</u>: At the time of diagnosis, there is no uncertainty in the real world. The patient either **has** cavity **(1)**, or **not (0)**.

Acting Under Uncertainty

One confusing point: At the time of diagnosis, there is no uncertainty in the real world. The patient either **has** cavity **(1)**, or **does not (0)**.

Probability statements are made with respect to a **knowledge state**, *not* to the real world it self.

Acting Under Uncertainty

Probability statements are made with respect to a **knowledge state**, *not* to the real world it self.

Example:

- Let’s say statistics say that 80% of patients *with* toothache have a cavity (Probability is 0.8)

Acting Under Uncertainty

Probability statements are made with respect to a **knowledge state**, *not* to the real world it self.

Example:

- ➔ Let's say statistics say that 80% of patients *with* toothache have a cavity (Probability is 0.8)
  - ◆ **A:** The probability that the patient has a cavity, given that she has a toothache, is 0.8.
  - ◆ **B:** The probability that the patient has a cavity, given that she has a toothache *and* a history of gum disease, is 0.4.
  - ◆ **C:** The probability that the patient has a cavity, given all we know, is 0

Uncertainty and Rational Decisions

- Consider A_{90}, and suppose it gives us a 97% chance of catching our flight.

Uncertainty and Rational Decisions

- Consider A_{90}, and suppose it gives us a 97% chance of catching our flight.
    - Rational? Not necessarily.

Uncertainty and Rational Decisions

- Consider A_{90}, and suppose it gives us a 97% chance of catching our flight.
  - Rational? Not necessarily.
- Consider A_{1440} (Leaving 24 hours before the flight)

Uncertainty and Rational Decisions

- Consider A_{90}, and suppose it gives us a 97% chance of catching our flight.
  - Rational? Not necessarily.
- Consider A_{1440} (Leaving 24 hours before the flight)
  - Not necessarily a good choice because very long wait and possibly very bad diet of airport food.

Uncertainty and Rational Decisions

- To make such choices, an agent must first have **preferences**, between possible **outcomes** of the plans.

Uncertainty and Rational Decisions

- To make such choices, an agent must first have **preferences**, between possible **outcomes** of the plans.
- We use the utility theory to represent and reason with “preference”

Uncertainty and Rational Decisions

- To make such choices, an agent must first have **preferences**, between possible **outcomes** of the plans.
- We use the utility theory to represent and reason with “preference”

Decision Theory = Probability Theory + Utility Theory

Uncertainty and Rational Decisions

- **Preference**: options, choices, what is more preferred.
- **Outcome**: Completely specified state, including such factors as whether the agent arrives on time and the length of the wait at the airport.
- **Utility Theory**: “The quality of being useful” - theory says that every state has a degree of usefulness, or utility, to an agent and that the agent will prefer states with higher utility.

Uncertainty and Rational Decisions

Maximum Expected Utility (MEU): *an agent is rational iff it chooses an action that yields the highest expected utility, averaged over all the possible outcomes of the action*.

Basic Probability Notation

Basic Probability Notation

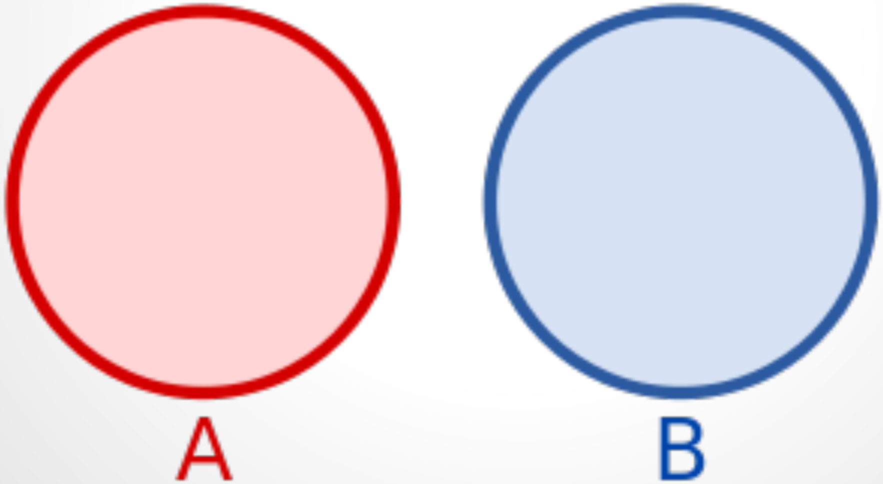
- **Sample Space**: in probability theory, the set of all possible worlds. (all possible outcomes)
- **Notation**:
  - **Ω** - ***Sample Space***
  - **ω** - *An **element** in the sample space*
  - **ϕ** - *An **event** or a **proposition***
    - An event (ϕ) is a subset of sample space (Ω): $\phi \subseteq \Omega$
    - Example: Two Dice adding up to 11 is an *event*
      - ϕ = { (5,6) , (6,5) }

Probability Model

- A probability model associates a numerical probability $P(\omega)$ with each possible world:
  - Every possible world must have a probability between 0 and 1
  - Total probability of the set of all possible worlds is 1
- When we say P(H) in a coin flip we refer to the probability of heads in a coin flip.
- P(H) = 0.5

Mutually Exclusive

- **Mutually Exclusive**: No two events have the same outcome
    - Mutually Exclusive: $A \cap B = \emptyset$



Exhaustive

- **Exhaustive**: at least one of the events must occur
    - Exhaustive: $A \cup B = \Omega$

Unconditional Probability

- Unconditional Probability is when you don't consider any other information
- Example:
  - Two dice: Blue, Red
  - You only consider the red die

Conditional Probability

- In Conditional Probability we have evidence (extra information) already revealed for example when rolling two dice if we know that the first die is a 6 then we know that the sum of the two die cannot be 5
- Conditional Probability: $P(A \mid B) = P(A \cap B) / P(B)$
- Product Rule: $P(A \land B) = P(A \mid B)\ P(B)$

Conditional Probability Example

- Events
  - A = {6}
  - B = {5}
  - $P(A \cap B) = 0$
  - $P(B) = \frac{1}{6}$
  - $P(A \mid B) = P(A \cap B) / P(B)$
  - $P(A \mid B) = 0 / \frac{1}{6}$
  - $P(A \mid B) = 0$

Random Variables

- Variables in Probability are called **Random Variables** and begin with an uppercase letter.
- Every Random Variable has a **domain** - a set of possible values that it can take. Domains usually start with a lowercase letter.
- For example, lets say we have the random variable *Total* that calculates the sum of two dice:
  - Then the domain is the set {2, …, 12} and Total(2) = 1/36
- A boolean random variable has the domain {True, False}

Probability Distribution

- A probability distribution is when we want to talk about all the possible values of a random variable. Usually indicated by a bold **P**.
- A **Discrete Random Variable** is a random variable that takes a finite number of distinct values
- **P**(Faircoin) = <0.5, 0.5>

Probability Density Function

- A **Continuous Random Variable** is a random variable that takes an infinite number of distinct values.
- Example:
  - **P**(NoonTemp = x) = $\text{Uniform}_{[18C,\ 26C]}(x)$ expresses that the temperature at noon is distributed uniformly between 18 and 26 degrees.
  - This is called a **probability density function**.

Joint Distribution

- The probabilities of all combinations of the values of the random variables
- Full joint probability distribution

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	*0.016*	0.064	0.144	0.576

Figure 13.3 A full joint distribu ion for the *Toothache, Cavity, Catch* world.

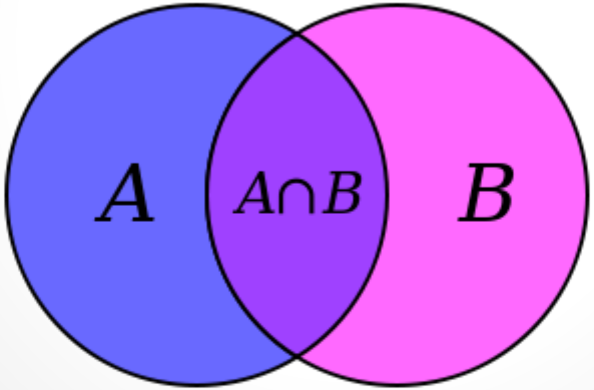
Basic Axioms of Probability

$$0 \le P(\omega) \le 1 \; for\ every\ \omega \text{ and } \sum_{\omega \in \Omega} P(\omega) = 1$$

$$For\ any\ proposition\ \phi, P(\phi) = \sum_{\omega \in \phi} p(\omega)$$

Inclusion-exclusion principle

- $P(A \cup B) = P(A)+P(B) - P(A \cap B)$



Independence

- Consider the previous example:
  - **P**(Toothache,Catch,Cavity,Weather)

Independence

- Consider the previous example:
    - **P**(Toothache,Catch,Cavity,Weather)

      How are they related?

Independence

- Consider the previous example:
  - **P**(Toothache,Catch,Cavity,Weather)

    How are they related?

P(toothache,catch,cavity,cloudy) =

P(cloudy|toothache,catch,cavity) P(toothache,catch,cavity)

Independence

- Consider the previous example:
  - **P**(Toothache,Catch,Cavity,Weather)

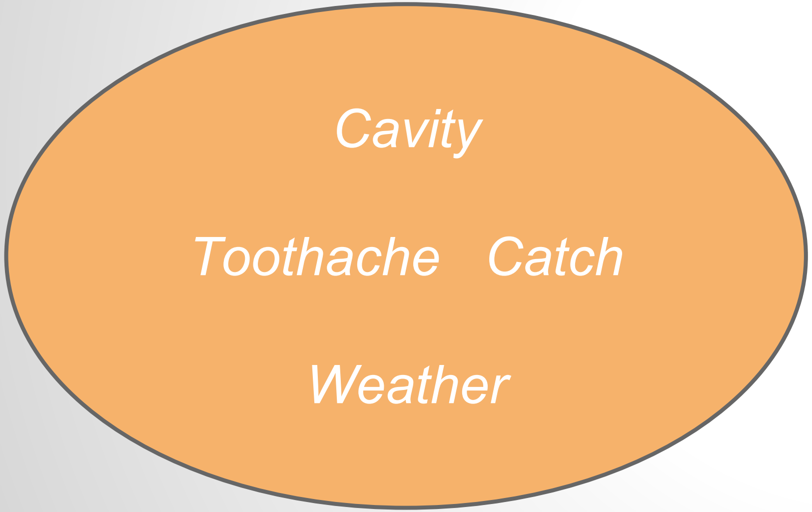
    How are they related?

P(toothache,catch,cavity,cloudy) =

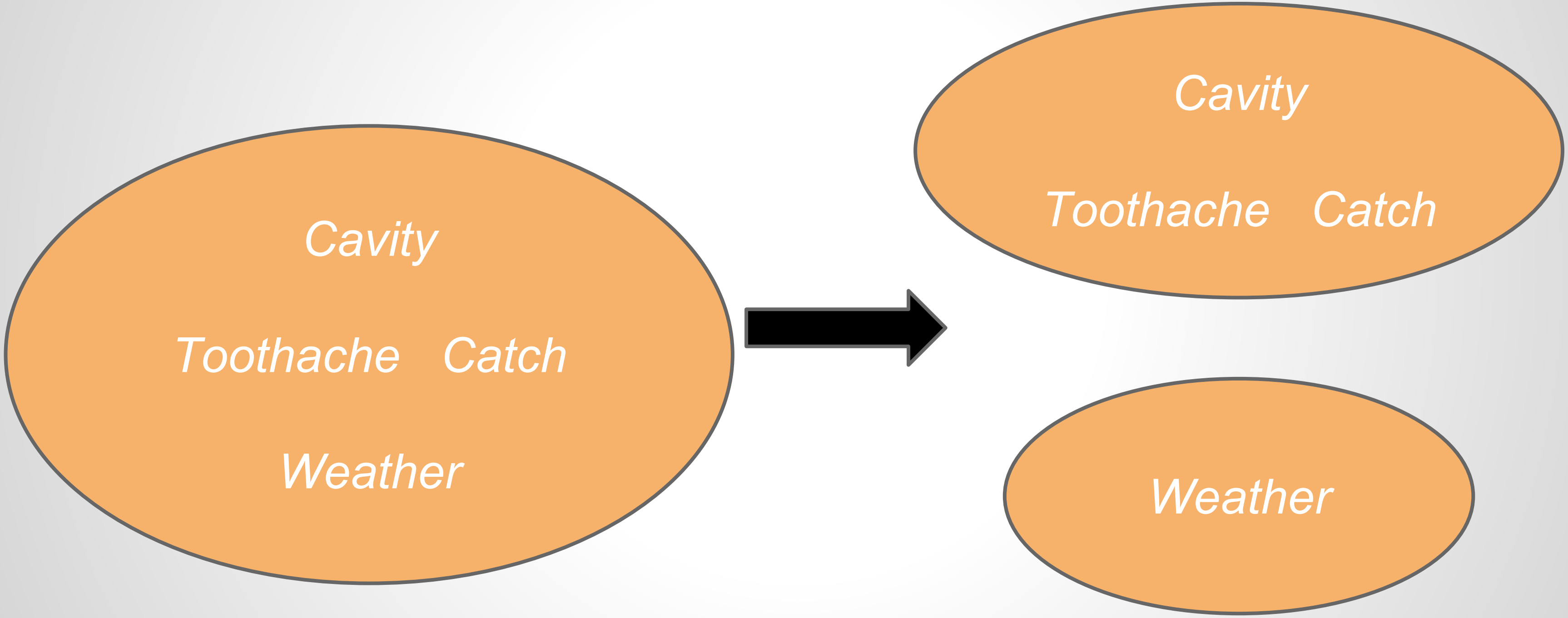
P(cloudy|toothache,catch,cavity) P(toothache,catch,cavity)

→ P(cloudy)

Independence



Independence



Independence

$$P(a \mid b) = P(a)$$

or

$$P(b \mid a) = P(b)$$

or

$$P(a \land b) = P(a)\,P(b)$$

Baye’s Rule

- Product Rule
  - $P(A \land B) = P(A \mid B)\, P(B)$
  - $P(A \land B) = P(B \mid A)\, P(A)$
- Baye’s Rule
  - $P(B \mid A) = P(A \mid B)\, P(B) / P(A)$

Applying Bayes Rule

$$P(s|m) = 0.7$$

$$P(m) = \frac{1}{50000}$$

$$P(s) = 0.01$$

$$P(m|s) = \frac{P(s|m)P(m)}{P(s)} = \frac{0.7 * 1/50000}{0.01} = 0.0014$$

Conditional Independence

- Conditional Independence of two variable X and Y, given a third variable Z, is
  - **P**(X, Y | Z) = **P**(X | Z) **P**(Y | Z)
  - **P**(X | Y, Z) = **P**(X | Z) and **P**(Y | X, Z) = **P**(Y | Z)

- **P**(Toothache, Catch, Cavity)
  - = **P**(Toothache, Catch | Cavity) **P**(Cavity)
  - = **P**(Toothache | Cavity) **P**(Catch | Cavity) **P**(Cavity)

Inference Using Full Joint Distribution

- Full Joint Probability Distribution: The joint distribution for all of the random variables.
  - Example: we have 3 random variables: Cavity, Toothache, and Weather. The table that shows all the probabilities of **P**(Cavity, Toothache, Weather) is the full joint probability distribution.

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	*0.016*	0.064	0.144	0.576

Figure 13.3 A full joint distribu ion for the *Toothache, Cavity, Catch* world.

Inference Using Full Joint Distribution

- Marginal Probability:
  - $P(Y) = \sum_{z \in Z} P(Y, z)$
    - $P(Cavity) = \sum_{z \in \{Catch, Toothache\}} P(Cavity, z)$
  - The sum of all probabilities for each possible value of the other variables of Z
  - Z is all the variables except Y.
  - $P(Y) = \sum_{z \in Z} P(Y \mid z) P(z)$ (conditioning)

Inference Using Full Joint Distribution

- Examples (from figure 13.3):
  - $P(cavity \vee toothache) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$
  - $P(cavity) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2$
  - $P(cavity | toothache) = \frac{P(cavity, toothache)}{P(toothache)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.6$

Inference Using Full Joint Distribution

- Normalization: is the process of scaling (by alpha scalar) some probabilities to add up to 1.
  - e.g.

$$P(Cavity \mid toothache) = \frac{1}{P(toothache)} \times P(Cavity,\ toothache) = \alpha P(Cavity,\ toothache)$$

$$= \alpha[P(Cavity, toothache, catch) + P(Cavity, toothache, \neg catch)]$$

$$= \alpha[< 0.108, 0.016 > \ + \ < 0.012, 0.064 >] = \alpha < 0.12, 0.08 >$$

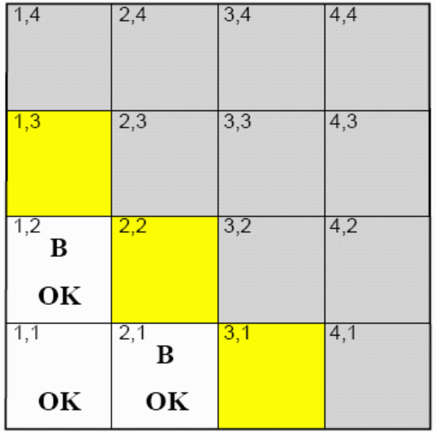
Let $\alpha = \frac{1}{0.12+0.08} = \frac{1}{0.2}$

$= \ < 0.6, 0.4 >$ (Check: 0.6+0.4=1)

  - Notice that: P(toothache)=0.108+0.012+0.016+0.064=0.2

Wumpus World Revisited

Wumpus World



$P_{ij} = true$ iff $[i, j]$ contains a pit

$B_{ij} = true$ iff $[i, j]$ is breezy
Include only $B_{1,1}, B_{1,2}, B_{2,1}$ in the probability model

Probabilistic Reasoning

Building network models to reason under uncertainty according to the laws of probability theory

Probabilistic Reasoning

- Chapter 14: Introduces a systematic way to represent independence and conditional independence relationships explicitly in the form of Bayesian Networks.
  - 14.1: Syntax of Bayesian Networks
  - 14.2: Semantics of Bayesian Networks

Representing Knowledge in Uncertain Domains

- Revisiting some definitions:
  - Independence: 2 events, A & B, are independents if they are unrelated to each other. That is, knowing one is true does not affect the other. (e.g. (A=Raining, B=Hungry))
    - i.e. $P(A|B)=P(A)$ or $P(B|A)=P(B)$
  - Conditional Independence: A & B are independents, given C is true. (i.e. $P(A|B,C)=P(A|C)$ or $P(B|A,C)=P(B|C)$)
  - Joint probability distribution: The probabilities of all combinations of the values of all Random Variables (events). (i.e. $P(X1,....,Xn)$)

Representing Knowledge in Uncertain Domains

- Network = Nodes + links (connections or directed arrows)
- Bayesian Network:
  - Represents any full joint probability distribution
  - Directed graph in which each node is is annotated with quantitative probability information.

Representing Knowledge in Uncertain Domains

- Specification of Bayesian Network:
  - Node = Random Variable
  - Directed Arrows
    - Directed Acyclic Graph (DAG)
    - $X \rightarrow Y$
      - X is a parent of Y
      - X has a direct influence on Y
      - X (cause) causes Y(effect)
  - Each node X_i has a conditional probability distribution $P(X_i | Parents(X_i))$

Representing Knowledge in Uncertain Domains

- Two types of Bayesian Networks:
  - Simple: Contains nodes and links (Helpful to view conditional relationships)
  - Typical: Simple Bayesian Network + conditional probability table (Good to represent joint probability distribution)

Representing Knowledge in Uncertain Domains

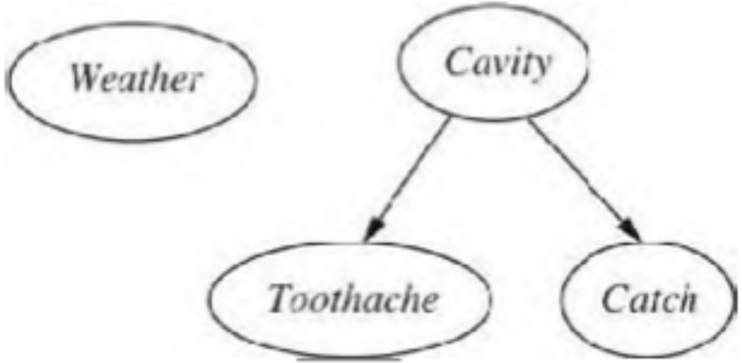


Figure 14.1 A simple Bayesian network in which *Weather* is independent of the other three variables and *Toothache* and *Catch*, are conditionally independent, given *Cavity*.

Burglary Example

Now consider the following example, which is just a little more complex. You have a new burglar alarm installed at home. it is fairly reliable at detecting a burglary, but also responds on occasion to minor earthquakes. (This example is due to Judea Pearl, a resident of Los Angeles—hence the acute interest in earthquakes.) You also have two neighbors, John and Mary, who have promised to call you at work when they hear the alarm. John nearly always calls when he hears the alarm, but sometimes confuses the telephone ringing with the alarm and calls then, too. Mary, on the other hand, likes rather loud music and often misses the alarm altogether. Given the evidence of who has or has not called, we would like to estimate the probability of a burglary.

Burglary Example

Now consider the following example, which is just a little more complex. You have a new burglar alarm installed at home. it is fairly reliable at detecting a burglary, but also responds on occasion to minor earthquakes. (This example is due to Judea Pearl, a resident of Los Angeles—hence the acute interest in earthquakes.) You also have two neighbors, John and Mary, who have promised to call you at work when they hear the alarm. John nearly always calls when he hears the alarm, but sometimes confuses the telephone ringing with the alarm and calls then, too. Mary, on the other hand, likes rather loud music and often misses the alarm altogether. Given the evidence of who has or has not called, we would like to estimate the probability of a burglary.

Burglary Example

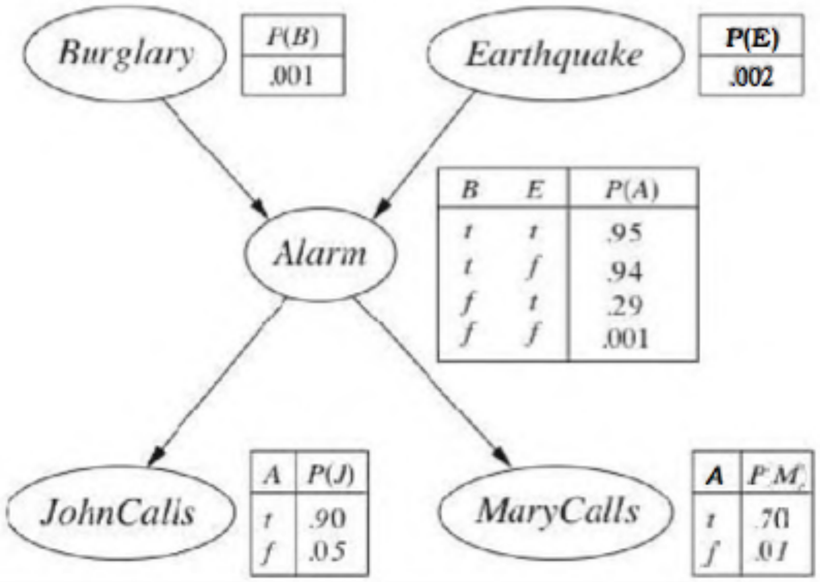


Figure 14.2 A typical Bayesian network, showing both the topology and the conditional probability tables (CPTs). In the CPTs, the letters *B*, *E*, *A*, *J*, and *M* stand for *Burglary*, *Earthquake*, *Alarm*, *JohnCalls*, and *MaryCalls*, respectively.

Semantics of Bayesian Network

- Two ways to view a Bayesian Network:
  - A representation of the Joint Probability Distribution
    - Helpful in understanding how to construct networks
  - Encoding a collection of Conditional Independence statements
    - Helpful in designing inference procedures
- The generic entry in the joint distribution is $P(X_1=x_1, .., X_n=x_n)$

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(x_i))$$

  - Example: n has sounded; neither a burglary nor an earthquake has occurred; both J & M call.

Constructing Bayesian Networks

Rewriting the entries in the joint distribution in terms of conditional probability (using product rule):

- $P(x_1, ..., x_n) = P(x_n \mid x_{n-1}, ..., x_1) P(x_{n-1}, ..., x_1)$

Convert each conjunctive probability to a conditional probability

- $P(x_1, ..., x_n) = P(x_n \mid x_{n-1}, ..., x_1) P(x_{n-1} \mid x_{n-2}, ..., x_1) P(x_{n-2}, .., x_1)$
- $P(x_1, ..., x_n) = P(x_n \mid x_{n-1}, ..., x_1) P(x_{n-1} \mid x_{n-2}, ..., x_1) ... P(x_2 \mid x_1) P(x_1)$
- $P(x_1, ..., x_n) = \prod_{i=1}^{n} P(x_i \mid x_{i-1}, ..., x_1)$ (Chain Rule)
  - Notice that $P(x_i \mid x_{i-1}, ..., x_1) = P(x_i \mid Parents(x_i))$

Constructing Bayesian Networks

1. Nodes: First determine the set of variables that are required to model the domain. Now order them, {Xi, *Xn}.* Any order will work, but the resulting network will be more compact if the variables are ordered such that causes precede effects.

2. Links: For *i* = 1 to n do;

- Choose, from {X1,... , *X* i -1} *,* a minimal set of parents for Xi, such that Equation (14.3) is satisfied.
- For each parent insert a link from the parent to Xi _
- CPTs: Write down the conditional probability table, P(Xi | Parents (Xi)).

Constructing Bayesian Networks

- Conditional Independence relations in Bayesian Networks:
  - A node is conditionally independent of its other predecessors, given its parents
  - Each variable is conditionally independent of its non-descendants, given its parents

References

- Russell, S. J., & Norvig, P. (2010).Artificial intelligence: a modern approach. Englewood Cliffs, N.J.: Prentice Hall.
- Images from:
  - http://www.ma.utexas.edu/users/rgrizzard/M316L_SP12/2dice.jpg
  - http://homepages.ius.edu/RWISMAN/C463/html/chapter13-27.gif