Review

- Knowledge Base
- Propositional Logic
- First Order Logic
  - Universal and Existential Quantifier
- Reduction of first-order inference to propositional inference
  - Universal and Existential Instantiation

Probability

Uncertainty and Vagueness

- Real world is not always exact and certain
  - Do the best with what is already known
  - Agent needs to draw conclusions when available information is uncertain, vague, or incomplete

Uncertainty and Vagueness

- Real world is not always exact and certain
  - Do the best with what is already known
  - Agent needs to draw conclusions when available information is uncertain, vague, or incomplete

- Sources of uncertainty
  - Unreliable data
    - Defective measurement device
  - Incomplete data
    - Only partial data available
  - Imprecise data/rules
    - Approximations of data
    - Rules for drawing conclusions may also be imprecise

Handling Uncertain Knowledge

- Consider dental diagnosis using first-order logic

$$\forall p \;\; Symptom(p, Toothache) \rightarrow Disease(p, Cavity)$$

  - Rule is wrong!
    - Not all patients with toothache have cavities

Handling Uncertain Knowledge

- Consider dental diagnosis using first-order logic

$$\forall p \;\; Symptom(p, Toothache) \rightarrow Disease(p, Cavity)$$

  – Rule is wrong!
    - Not all patients with toothache have cavities

  – Need to add an almost <u>unlimited</u> list of possible causes

$$\begin{aligned} \forall p \;\; Symptom(p, Toothache) \rightarrow {} & Disease(p, Cavity) \\ & \lor Disease(p, GumDisease) \\ & \lor Disease(p, ImpactedWisdom) \ldots \end{aligned}$$

Handling Uncertain Knowledge

- Use of first-order logic fails in diagnosis for three reasons
  - Laziness
  - Theoretical ignorance
  - Practical ignorance

Handling Uncertain Knowledge

- Use of first-order logic fails in diagnosis for three reasons
  - Laziness
    - Too much work to list all possible consequents
  - Theoretical ignorance
  - Practical ignorance

Handling Uncertain Knowledge

- Use of first-order logic fails in diagnosis for three reasons
  - <u>Laziness</u>
    - Too much work to list all possible consequents
  - <u>Theoretical ignorance</u>
    - No complete theory for domain available
  - <u>Practical ignorance</u>

Handling Uncertain Knowledge

- Use of first-order logic fails in diagnosis for three reasons
  - Laziness
    - Too much work to list all possible consequents
  - Theoretical ignorance
    - No complete theory for domain available
  - Practical ignorance
    - Even if have all rules, may be uncertain about a patient (all tests have not been run to get data)

Handling Uncertain Knowledge

- System's knowledge (Knowledge Base) can **only** provide a <u>degree of belief</u> about sentences
  - Sentences are in fact either true or false

Handling Uncertain Knowledge

- System's knowledge (Knowledge Base) can **only** provide a <u>degree of belief</u> about sentences
  - Sentences are in fact either true or false
- Probability theory
  - Deals with uncertainty
  - Assignment of probability depends on percepts received to date (evidence)
  - Assigns numerical degree of belief between 0-1

Handling Uncertain Knowledge

- System's knowledge (Knowledge Base) can **only** provide a degree of belief about sentences
  - Sentences are in fact either true or false
- Probability theory
  - Deals with uncertainty
  - Assignment of probability depends on percepts received to date (evidence)
  - Assigns numerical degree of belief between 0-1
    - Probability of 0 absolute belief that sentence if FALSE
    - Probability of 1 is absolute belief that sentence is TRUE
    - Probability of 0.8 (80%) that patient has cavity if has toothache

Basics of Probability

- Prior probability
- Conditional probability
- Bayes rule

Basic Concepts of Probability

- Prior probability
  - *A priori* initial information: $P(A)$
    - "Probability of event A occurring"

Basic Concepts of Probability

- Prior probability
  - *A priori* initial information: $P(A)$
    - "Probability of event A occurring"
  - Independent of any evidence

Basic Concepts of Probability

- Joint probability: $P(A \wedge B)$
  - "Probability of event A <u>and</u> B occurring together"
  - Also often written $P(A, B)$

Basic Concepts of Probability

- Conditional probability
  - Information based on additional evidence: $P(A \mid C)$
    - "Probability of A, given that I know C"

Prior Probability

Prior Probability

- Probability assessment <u>before</u> any evidence obtained
  - “Probability of event A” $\rightarrow$ $P(A)$
  - e.g., $P(Cavity) = 0.1$
    - In the absence of any other information, 10% chance of the patient having a cavity

Prior Probability

- Probability assessment <u>before</u> any evidence obtained
  - "Probability of event A" $\rightarrow$ $P(A)$
  - e.g., $P(Cavity) = 0.1$
    - In the absence of any other information, 10% chance of the patient having a cavity

- Random variable assignments of exclusive values

$$\left.\begin{aligned} &P(\mathbf{weather} = Sunny) = 0.7 \\ &P(\mathbf{weather} = Rainy) = 0.2 \\ &P(\mathbf{weather} = Cloudy) = 0.08 \\ &P(\mathbf{weather} = Snow) = 0.02 \end{aligned}\right\} \quad \text{Must add up to 1}$$

$$\mathbf{weather} \rightarrow \{Sunny, Rainy, Cloudy, Snow\}$$

Prior Probability

- Properties defining probability $P(A)$

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

$$P(\text{“}\mathit{certain\ event}\text{”}) = 1$$

$$P(A) + P(\neg A) = 1$$

$$P(\neg A) = 1 - P(A)$$

Prior Probability

- Properties defining probability $P(A)$

$$0 \le P(A) \le 1$$
$$P(\text{“}\mathit{certain\ event}\text{”}) = 1$$
$$P(A) + P(\neg A) = 1$$
$$P(\neg A) = 1 - P(A)$$

- Simple die examples
  - Each side of die labeled from 1 to 6
  - Let event A be a die stops with 1 showing on top: $P(A) = 1/6$
  - Let event ~A be a die stops with numbers other than a 1 showing on top

$$P(\neg A) = ?$$
$$\begin{aligned} P(\neg A) &= 1 - P(A) \\ &= 1 - 1/6 \\ &= 5/6 \end{aligned}$$

Joint Probability

- Joint probability ("and")

  $P(Cavity \wedge \neg Insured) = 0.06$

Prior & Joint Probability

- Joint probability ("and")

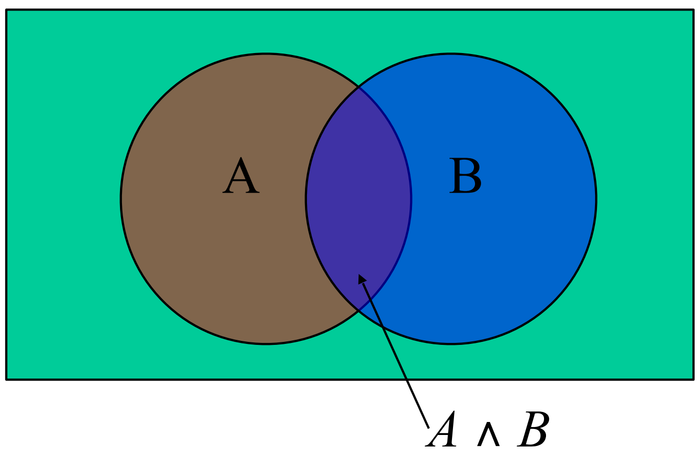
$$P(\mathit{Cavity} \land \neg \mathit{Insured}) = 0.06$$

- Marginalization

$$P(\mathit{Cavity}) = P(\mathit{Cavity} \land \mathit{Insured}) + P(\mathit{Cavity} \land \neg \mathit{Insured})$$

Probability axiom of a Disjunction (“or”)

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$



Conditional Probability

Conditional Probability

- Probability assessment <u>after</u> evidence is obtained ("posterior probability")
  - "Probability of event A given we know B"

    $\rightarrow P(A \mid B)$
  - e.g., $P(Cavity \mid Toothache) = 0.8$
    - If observe patient with toothache, 80% chance patient has cavity

Conditional Probability

- Probability assessment <u>after</u> evidence is obtained ("posterior probability")
  - "Probability of event A given we know B"
    $\rightarrow P(A \mid B)$
  - e.g., $P(Cavity \mid Toothache) = 0.8$
    - If observe patient with toothache, 80% chance patient has cavity
- Properties defining conditional probability $P(A|B)$

$$0 \leq P(A|B) \leq 1$$

$$P(A|B) + P(\neg A|B) = 1$$

$$P(\neg A|B) = 1 - P(A|B)$$

Conditional Probability

- Probability assessment after evidence is obtained ("posterior probability")
  - "Probability of event A given we know B"
    $\rightarrow P(A \mid B)$
  - e.g., $P(Cavity \mid Toothache) = 0.8$
    - If observe patient with toothache, 80% chance patient has cavity
- Properties defining conditional probability $P(A|B)$

$$0 \leq P(A|B) \leq 1$$

$$P(A|B) + P(\neg A|B) = 1$$
$$P(\neg A|B) = 1 - P(A|B)$$

$$P(A|\neg B) + P(\neg A|\neg B) = 1$$
$$P(\neg A|\neg B) = 1 - P(A|\neg B)$$

Conditional Probability

- Probability assessment <u>after</u> evidence is obtained ("posterior probability")
  - "Probability of event A given we know B"
    - $\rightarrow P(A \mid B)$
  - e.g., $P(Cavity \mid Toothache) = 0.8$
    - If observe patient with toothache, 80% chance patient has cavity
- Properties defining conditional probability $P(A|B)$

$$0 \leq P(A|B) \leq 1$$

$$P(A|B) + P(\neg A|B) = 1$$
$$P(\neg A|B) = 1 - P(A|B)$$

$$P(A|\neg B) + P(\neg A|\neg B) = 1$$
$$P(\neg A|\neg B) = 1 - P(A|\neg B)$$

<u>No relation</u> between $P(A|B)$ and $P(A|\neg B)$

Conditional Probability

- Conditional probability in terms of unconditionals

  $P(A \wedge B) = P(A \mid B) \cdot P(B)$ [Product or multiply rule]

  thus

  $P(A \mid B) = P(A \wedge B) / P(B)$

Conditional Probability

- Conditional probability in terms of unconditionals

  $P(A \wedge B) = P(A \mid B) \cdot P(B)$ [Product or multiply rule]

  thus

  $P(A \mid B) = P(A \wedge B) \,/\, P(B)$

  Note: If B <u>independent</u> of A, changes to B do not affect A:

  $P(A \mid B) = P(A \mid \neg B) = P(A)$

  $P(A \wedge B) = P(A \mid B) \cdot P(B) = P(A)P(B)$

Conditional Probability

- As soon as get more information C, must compute new conditional

  $$P(A \mid B \wedge \mathrm{C})$$

Conditional Probability

- As soon as get more information C, must compute new conditional

  $$P(A \mid B \wedge \mathrm{C})$$

- If A and B are <u>conditionally independent</u> given evidence C

  $$P(A \mid B \wedge \mathrm{C}) = P(A \mid C)$$

Inference from Joint Probabilities

- Table of 2 propositions: *Cavity* and *Toothache*

	Toothache	$\neg$ *Toothache*
Cavity	0.04	0.06
$\neg$ *Cavity*	0.01	0.89

$P(Cavity \wedge \neg Toothache)$

$P(Toothache)$

$$P(Cavity \vee Toothache) = P(Cavity) + P(Toothache) - \mathrm{P}(\mathrm{C} \wedge \mathrm{T})$$
$$= (0.04 + 0.06) + (0.04 + 0.01) - 0.04 = 0.11$$

$$P(Cavity \mid Toothache) = P(Cavity \wedge Toothache)/P(Toothache)$$
$$= 0.04 / (0.04 + 0.01) = 0.80$$

Bayes Rule

Bayes Rule

- Typical problem is to evaluate a hypothesis H given evidence/data E:
    - $P(H \mid E)$

Bayes Rule

- Typical problem is to evaluate a hypothesis H given evidence/data E:
  - $P(H \mid E)$
- Recall our earlier product rule for a conditionals

$$P(H \mid E) = P(H \wedge E) \,/\, P(E)$$

But can be hard to get or understand $P(H \wedge E)$

Bayes Rule

- Typical problem is to evaluate a hypothesis H given evidence/data E:
  - $P(H \mid E)$
- Recall our earlier product rule for a conditionals

$$P(H \mid E) = P(H \wedge E) \,/\, P(E)$$

But can be hard to get or understand $P(H \wedge E)$

- Recall product rule (multiplicative law)

$$P(A \wedge B) = P(A \mid B) \cdot P(B)$$
$$P(B \wedge A) = P(B \mid A) \cdot P(A)$$

Bayes Rule

- Typical problem is to evaluate a hypothesis H given evidence/data E:
  - $P(H \mid E)$
- Recall our earlier product rule for a conditionals

$$P(H \mid E) = P(H \wedge E) \,/\, P(E)$$

But can be hard to get or understand $P(H \wedge E)$

- Recall product rule (multiplicative law)

$$P(A \wedge B) = P(A \mid B) \cdot P(B)$$
$$P(B \wedge A) = P(B \mid A) \cdot P(A)$$

[*equate right-hand sides* $[P(A \wedge B) = P(B \wedge A)]$]

Bayes Rule

- Typical problem is to evaluate a hypothesis H given evidence/data E:
    - $P(H \mid E)$
- Recall our earlier product rule for a conditionals

$$P(H \mid E) = P(H \wedge E) \,/\, P(E)$$

But can be hard to get or understand $P(H \wedge E)$

- Recall product rule (multiplicative law)

$$P(A \wedge B) = P(A \mid B) \cdot P(B)$$
$$P(B \wedge A) = P(B \mid A) \cdot P(A)$$

[*equate right-hand sides* $[P(A \wedge B) = P(B \wedge A)]$]

$$P(A \mid B) \cdot P(B) = P(B \mid A) \cdot P(A)$$
$$P(A \mid B) = P(B \mid A) \cdot P(A) \,/\, P(B)$$
$$P(H \mid E) = P(E \mid H) \cdot P(H) \,/\, P(E) \longleftarrow \text{Bayes' rule}$$

Bayes Rule

- Other variations of denominator

BAYES: $P(H \mid E) = P(E \mid H)P(H) \,/\, P(E)$

$$P(E) = P(E \wedge H) + P(E \wedge \neg H)$$
$$= P(E \mid H)P(H) + P(E \mid \neg H)P(\neg H)$$
$$= \Sigma_{\mathrm{i}}\, P(E \mid H_i)P(H_i)$$

Bayes Rule

- Other variations of denominator

BAYES: $P(H \mid E) = P(E \mid H)P(H) \,/\, P(E)$

$$P(E) = P(E \wedge H) + P(E \wedge \neg H)$$
$$= P(E \mid H)P(H) + P(E \mid \neg H)P(\neg H)$$
$$= \Sigma_{\mathrm{i}}\, P(E \mid H_i)P(H_i)$$

"sum/integrate over all possible hypotheses"

Bayes Rule

- Other variations of denominator

BAYES: $P(H \mid E) = P(E \mid H)P(H) \;/\; P(E)$

$$
\begin{aligned}
P(E) &= P(E \wedge H) + P(E \wedge \neg H) \\
&= P(E \mid H)P(H) + P(E \mid \neg H)P(\neg H) \\
&= \Sigma_{\mathrm{i}}\, P(E \mid H_i)P(H_i)
\end{aligned}
$$

"sum/integrate over all possible hypotheses"

$$
P(H \mid E) = \frac{P(E \mid H)P(H)}{\Sigma_{\mathrm{i}}\, P(E \mid H_i)P(H_i)}
$$

Summary

- Decisions not always exact and certain
  - Agent needs to draw conclusions when available information is uncertain
- Probability laws
  - Prior probability
    - A priori initial information, independent of experience
  - Joint probability
  - Conditional probability
    - Information based on additional evidence
  - Bayes rule
    - Updates belief measures in response to evidence