COMPSCI 366 S1 C 2006
Foundations of Artificial Intelligence

—Reasoning under Uncertainty—

Hans W. Guesgen
Computer Science Department

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

Why Reasoning under Uncertainty?

- Available knowledge is often incomplete or inexact (for example, weather prediction and medical diagnosis).
- Humans have ways of drawing inferences from incomplete, inexact, or uncertain knowledge.
- Humans make and use generalizations and approximations, which are often subject to error.

The Notion of Certainty

- Unlikely but possible.
- I don't know and I don't care.
- Sure, impossible, maybe, . . .
- I'll give you 10 to 1 it's true.
- . . .

Laplace's Formula

$$\text{Probability} = \frac{\text{Number of desired outcomes}}{\text{Total number of outcomes}}$$

Axioms of Probability

Additive law:

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

Multiplicative law:

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$$

Example

- $H =$ "John has malaria."
- $E =$ "John has a high fever."
- $P(H) =$ probability that a person has malaria.
- $P(E|H) =$ probability that a person has a high fever, given that he or she has malaria.
- $P(E|\neg H) =$ probability that a person has a high fever, given that he or she does not have malaria.

Bayes's Rule

- The probability that John has malaria, given that he has a high fever, is equal to the ratio of the probability that he has both the fever and malaria to the probability that he has a fever regardless of whether or not he has malaria:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

- The probability of having a high fever is computed as the sum of the conditional probabilities of having the fever given malaria or given not malaria, weighted by the probability of malaria and not malaria, respectively:

$$P(E) = P(E|H)P(H) + P(E|\neg H)P(\neg H)$$

Problem with Bayes's Rule

- If we always had accurate general knowledge about all the evidence, then building an inference system based on Bayes's rule would be simple.
- Unfortunately, we usually do not have accurate knowledge of the conditional probabilities of evidence (e.g., sets of symptoms) given the hidden truth (e.g., the state of health).

Probabilistic Inference Networks

- Model for general decision making in a computationally practical yet mathematically meaningful way.
- Good at handling information processing tasks with the following characteristics:
  1. Pieces of information are available at various levels of certainty and completeness.
  2. There is a need for optimal or nearly optimal decisions.
  3. There may be a need to justify the arguments in favor of the leading alternative choices.
  4. General rules of inference (either based on scientific theory or simply heuristic) are known or can be found for the problem.

Application Areas

- Medical diagnosis
- Fault diagnosis in machines and software
- Minerals prospecting
- Criminal investigations
- Military strategy formulation
- Marketing strategy and investment
- Decision making in design processes

Inference Rules

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \oplus B$
Propositional logic						
F	F	T	F	F	T	F
F	T	T	F	T	T	T
T	F	F	F	T	F	T
T	T	F	T	T	T	F
Possibilistic logic						
a	b	$1-a$	$\min(a,b)$	$\max(a,b)$	$\max(1-a,b)$	$\mathsf{xor}(a,b)$
Probabilistic logic						
a	b	$1-a$	ab	$a+b-ab$	$1-a+ab$	$\mathsf{Xor}(a,b)$

$$\mathsf{xor}(a,b) = \max[\min(a,1-b), \min(1-a,b)]$$
$$\mathsf{Xor}(a,b) = a + b - 3ab + a^2b + ab^2 - a^2b^2$$

The Automobile Repair Example

- S_1: There is a clanking sound in the engine.
- S_2: The car is low on pickup.
- S_3: The engine has trouble starting.
- S_4: Parts are difficult to obtain for this make.

- C_1: The repair estimate is over $250.

- H_1: A connecting rod is thrown in the engine.
- H_2: A wrist pin is loose.
- H_3: The car is out of tune.
- H_4: The engine needs rebuilding.
- H_5: The engine needs a tune-up.

The Automobile Repair Example (cont'd)

Symptoms			$P(S\|H_1)$	$P(S\|H_2)$	$P(S\|H_3)$	$P(S)$
S_1	S_2	S_3	$P(H_1) = 0.0001$	$P(H_2) = 0.0002$	$P(H_3) = 0.1$	
F	F	F	0.001	0.2	0.2	0.4405
F	F	T	0.003	0.1	0.2	0.25
F	T	F	0.006	0.1	0.2	0.109
F	T	T	0.15	0.1	0.396	0.2
T	F	F	0.04	0.125	0.001	0.0001
T	F	T	0.06	0.125	0.001	0.0001
T	T	F	0.11	0.125	0.001	0.0001
T	T	T	0.63	0.125	0.001	0.0001

The Automobile Repair Example (cont'd)

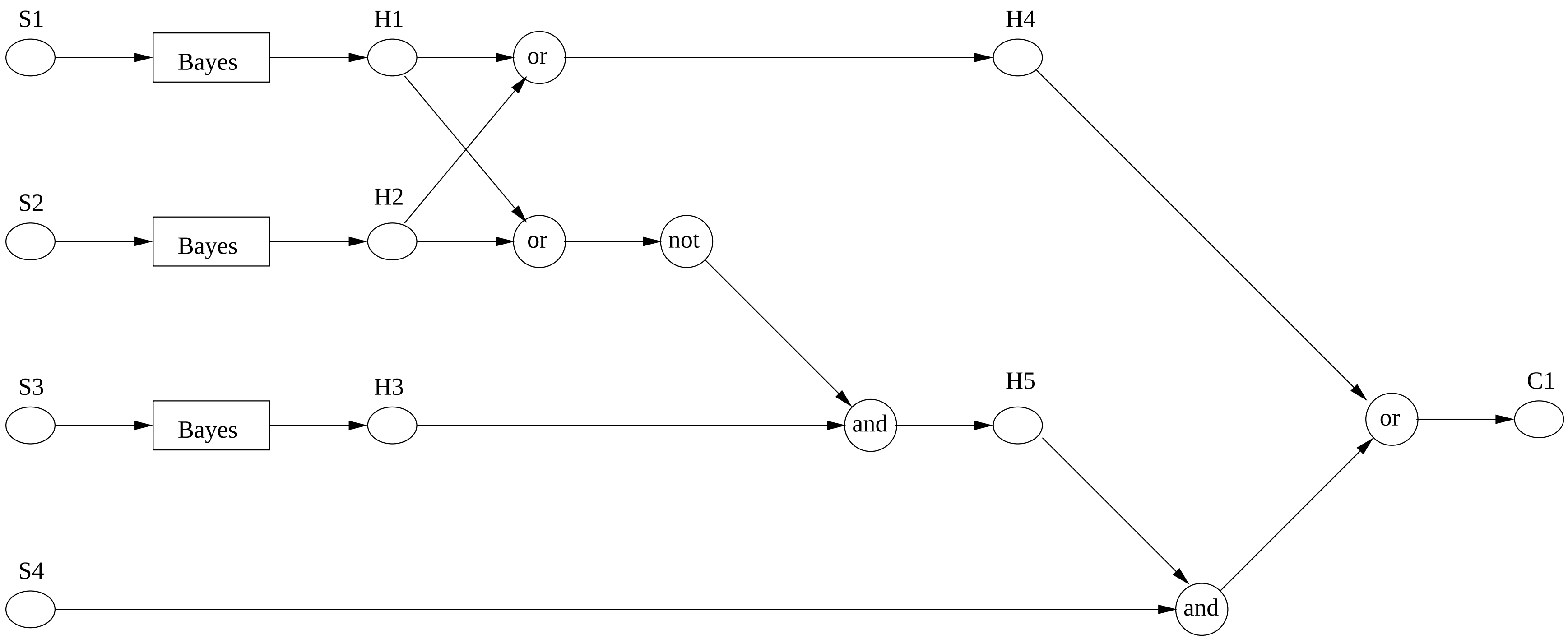
$$H_4 = H_1 \vee H_2$$
$$P(H_4|S) = \max[P(H_1|S), P(H_2|S)]$$

$$H_5 = \neg(H_1 \vee H_2) \wedge H_3$$
$$P(H_5|S) = \min\{1 - \max[P(H_1|S), P(H_2|S)], P(H_3|S)\}$$

$$C_1 = H_4 \vee (H_5 \wedge S_4)$$
$$P(C_1|S) = \max[P(H_4|S), \min(P(H_5|S), v)]$$

where $v = 1$, if S_4 is true; 0 otherwise

Probabilistic Inference Network



How to Design an Inference Network

1. Determine the relevant inputs (i.e., the set of possible evidence or symptoms).

2. Determine the states of nature or decision alternatives.

3. Determine the intermediate assertions that may be useful in the inference network.

4. Formulate the inference links.

5. Tune the probabilities and/or fuzzy inference functions.

Counterintuitive Measures of Belief

- The use of Bayes's rule for manipulating measures of belief is often regarded as inappropriate.
- Belief measures should not have to behave as probabilities.
- Example:
  - Propositon A: *Acme computers are intelligent.*
  - Sam doesn't know what a computer is.
  - $B(A)$ is Sam's degree of belief that A is true;
    $B(\neg A)$ is Sam's degree of belief that A isn't true.
  - $B(A)$ isn't necessarily equal to $1 - B(\neg A)$;
    it is more reasonable to assign $B(A) = B(\neg A) = 0$.

Dempster-Shafer Calculus

- System for manipulating degrees of belief that is more general than the Bayesian approach.
- Does not require the assumption that $B(A) + B(\neg A) = 1$.
- Somewhat more complicated than the Bayesian approach.

Universe and Basic Probability Assigment

- Let U be a (usually) finite set called the universe or frame of discernment.
- Then the set of all propositions is $\mathcal{P}(U)$, the power set of U.
- It contains the set U, which may be considered certainly true, and the proposition $\emptyset$, which may be considered certainly false.
- A function m with the following properties is called a basic probability assignment:
  1. $m(\emptyset) = 0$
  2. $\sum_{A \subseteq U} m(A) = 1$

Belief, Doubt, and Plausibility

$$\begin{aligned}
\textbf{Belief}(A) &= \textstyle\sum_{B \subseteq A} m(B) \\
\textbf{Doubt}(A) &= \textbf{Belief}(\neg A) \\
\textbf{Plausibility}(A) &= 1 - \textbf{Doubt}(A)
\end{aligned}$$

$$\begin{aligned}
\textbf{Belief}(\emptyset) &= 0 & \qquad \textbf{Plausibility}(\emptyset) &= 0 \\
\textbf{Belief}(U) &= 1 & \qquad \textbf{Plausibility}(U) &= 1
\end{aligned}$$

$$\begin{aligned}
\textbf{Plausibility}(A) &\geq \textbf{Belief}(A) \\
\textbf{Belief}(A) + \textbf{Belief}(\neg A) &\leq 1 \\
\textbf{Plausibility}(A) + \textbf{Plausibility}(\neg A) &\geq 1
\end{aligned}$$

For $B \subseteq A$:

$$\begin{aligned}
\textbf{Belief}(B) &\leq \textbf{Belief}(A) \\
\textbf{Plausibility}(B) &\leq \textbf{Plausibility}(A)
\end{aligned}$$

Dempster's Rule of Combination

- Orthogonal sum of basic probability assignments m_1 and m_2:
  - If $A \neq \emptyset$:

$$[m_1 \oplus m_2](A) = \frac{\sum_{X \cap Y = A} m_1(X) m_2(Y)}{1 - \sum_{X \cap Y = \emptyset} m_1(X) m_2(Y)}$$

  - Otherwise:

$$[m_1 \oplus m_2](\emptyset) = 0$$

- "Weight of conflict" must not be equal to 1:

$$\sum_{X \cap Y = \emptyset} m_1(X) m_2(Y) \neq 1$$

Example

- Two pieces of uncertain evidence:
  1. The temperature today is below freezing.
  2. The barometric pressure is falling, i.e., a storm is likely.
- Basic probability assignments for snow, rain, and dry weather:

	$\emptyset$	$\{S\}$	$\{R\}$	$\{D\}$	$\{S, R\}$	$\{S, D\}$	$\{R, D\}$	$\{S, R, D\}$
m_{freeze}	0.0	0.2	0.1	0.1	0.2	0.1	0.1	0.2
m_{storm}	0.0	0.1	0.2	0.1	0.3	0.1	0.1	0.1
m_{both}	0.0	0.282	0.282	0.128	0.180	0.051	0.051	0.026