MEASI
Institute *of*
Information Technology

Ms.Rafath Zahra

Representing knowledge using rules

Knowledge can be represented in two ways:

- declarative
- procedural

Declarative knowledge involves understanding facts or information, such as knowing "the sun rises in the east." This type of knowledge is static, simply describing how things are, such as concepts, propositions, or relationships between different entities.

procedural knowledge involves understanding how to perform tasks or procedures, such as "how to make coffee." This type of knowledge is dynamic, focusing on the steps required to achieve a specific goal.

When we use knowledge in the context of rules, we can employ methods like **forward chaining**.

This method begins with known facts and applies inference rules to derive new information, step by step, until a specific conclusion or action is reached. For example, if we have the fact "The sky is cloudy" and a rule "If the sky is cloudy, then it might rain," forward chaining would lead us to conclude that it might rain.

Declarative knowledge provides the facts, while procedural knowledge helps in utilizing those facts to perform actions or solve problems through methods like forward chaining.

Ms.Rafath Zahra

Representing knowledge using rules

.

Knowledge Base: This is the core of the system, containing a collection of rules and facts. The rules are typically in the form of "if-then" statements, where the "if" part represents a condition and the "then" part represents the conclusion or action to be taken if the condition is met.

Inference Engine: This is the processing unit that applies the rules in the knowledge base to the known facts or data. The inference engine can operate in two main modes:

Forward Chaining: This is a data-driven approach where the inference engine starts with known facts and applies rules to derive new facts or conclusions, moving from the specific data to a general conclusion.

Backward Chaining: This is a goal-driven approach where the inference engine starts with a goal or hypothesis and works backward to see if the known facts and rules support the goal.

Working Memory: This is where the current known facts are stored. As the inference engine applies rules and deduces new facts, these facts are added to the working memory.

Rule-Based Deduction System

Example:
Consider a medical diagnosis system as an example of a rule-based deduction system:

Rule 1: If the patient has a fever and a cough, then suspect flu.
Rule 2: If the patient has a rash and a fever, then suspect measles.
Rule 3: If the patient has flu, recommend rest and fluids.

Forward Chaining: **Scenario:** A patient comes in with symptoms: fever and cough.

The system starts with the facts: the patient has a fever and a cough.
Apply Rule 1: The system checks if any rule can be applied to these facts.
Rule 1: "If the patient has a fever and a cough, then suspect flu."
Since both conditions (fever and cough) match, the system concludes: suspect flu.
Add New Fact: The system adds "suspect flu" to the working memory as a new fact.
Apply Rule 3: Now the system checks if any rule can be applied to this new fact.
Rule 3: "If the patient has flu, recommend rest and fluids."
Since "suspect flu" is in the working memory, the system concludes: recommend rest and fluids.

Rule-Based Deduction System

- **Backward Chaining: Scenario:** A doctor wants to know if they should recommend rest and fluids.
- **Start with the Goal:** The system starts with the goal or query: Should we recommend rest and fluids?
- **Look for a Matching Rule:** The system looks for a rule that has this goal in the "then" part.
- **Rule 3:** "If the patient has flu, recommend rest and fluids."
- The system sees that to recommend rest and fluids, it needs to confirm whether the patient has the flu.
- **Set a Sub-Goal:** The system now needs to confirm whether the patient has the flu.
- To do this, it looks for a rule that could lead to the conclusion that the patient has flu.
- **Rule 1:** "If the patient has a fever and a cough, then suspect flu."
- The system sets a sub-goal to check if the patient has a fever and a cough.
- **Check Initial Facts:** The system checks the initial facts: does the patient have a fever and a cough?
- If both facts are true, then Rule 1 is satisfied, so the system can confirm suspect flu.
- **Achieve the Initial Goal:** With the flu now confirmed, Rule 3 can be applied, leading to the conclusion to recommend rest and fluids.

Ms.Rafath Zahra

Reasoning under uncertainty

Reasoning under uncertainty involves making decisions when outcomes aren't certain.

We use probability to measure the likelihood of events, which helps us make informed choices.

Key Terms:

Probability: Ranges from 0 (impossible) to 1 (certain). For example, flipping a coin has a probability of 0.5 for heads.

Conditional Probability: Updates beliefs based on new information, e.g., the likelihood of having a disease given a positive test result.

Bayes' Theorem: A formula to update the probability of an event based on new evidence.

Ms.Rafath Zahra

Decision Making

.

When making decisions where the outcome is uncertain, we need a method to choose the best option. Here's how it works:

Basic Ideas:

- Probability: This tells us how likely something is to happen.
- Utility: This tells us how much we like or value an outcome.
- Expected Utility: This combines both probability and utility to show us the average value of an action considering all possible outcomes.

How to Decide:

- Understand What You Know: Figure out the likelihood of different outcomes.
- Predict What Will Happen: Think about what will happen with each option you have.
- Pick the Best Option: Choose the option that gives you the highest average value, considering both how likely it is and how much you value the outcome.
- By using this method, you can make the best choices even when you're not sure what will happen.

Ms.Rafath Zahra

Key Terms

Random Variables

A random variable is a way to represent a part of the world whose outcome is uncertain. It can take on different values based on chance.

Types:

Boolean Random Variables: These can be either true or false. Example: "Cavity" could be true (you have a cavity) or false (you don’t).

Discrete Random Variables: These take on a set of distinct values. Example: "Weather" could be sunny, rainy, cloudy, or snowy.

Continuous Random Variables: These can take on any value within a range. Example: The exact temperature tomorrow can be any number within a range.

Key Terms

Atomic Events

Atomic events describe every possible combination of values for all the random variables.

For example, if you have two variables, "Cavity" and "Toothache," an atomic event could be "Cavity = true and Toothache = false."

Properties:

- Mutually Exclusive: Two events are mutually exclusive if they cannot both happen at the same time.
- Example: Imagine you're rolling a six-sided die. The events "rolling a 3" and "rolling a 5" are mutually exclusive because you cannot roll a 3 and a 5 simultaneously on a single roll

- Exhaustive: A set of events is exhaustive if they cover all possible outcomes, meaning one of them must happen.
- Example: Using the die again, the events "rolling a 1," "rolling a 2," "rolling a 3," "rolling a 4," "rolling a 5," and "rolling a 6" are exhaustive because one of these outcomes must occur every time you roll the die. There are no other possible outcomes.

Ms.Rafath Zahra

Key Terms

Atomic Events
Atomic events describe every possible combination of values for all the random variables.

For example, if you have two variables, "Cavity" and "Toothache," an atomic event could be "Cavity = true and Toothache = false."

Properties:

- Mutually Exclusive: Two events are mutually exclusive if they cannot both happen at the same time.
- Example: Imagine you're rolling a six-sided die. The events "rolling a 3" and "rolling a 5" are mutually exclusive because you cannot roll a 3 and a 5 simultaneously on a single roll

- Exhaustive: A set of events is exhaustive if they cover all possible outcomes, meaning one of them must happen.
- Example: Using the die again, the events "rolling a 1," "rolling a 2," "rolling a 3," "rolling a 4," "rolling a 5," and "rolling a 6" are exhaustive because one of these outcomes must occur every time you roll the die. There are no other possible outcomes.

Ms.Rafath Zahra

Key Terms

Probability Distributions
Probability distributions tell us the likelihood of different values for a random variable.

Discrete Probability Distribution: For discrete variables, it lists the probability for each possible value. For example, for the "Weather" variable, it could be:
Sunny: 70%
Rainy: 20%
Cloudy: 8%
Snowy: 2%

Joint Probability Distribution: Shows the probability of different combinations of values for multiple variables.

Example: For variables like "Weather" (Sunny, Rainy) and "Cold" (Yes, No), the joint probability distribution shows the probability of each combination, like "Sunny and Cold = Yes."

Full Joint Probability Distribution: Covers every possible combination of values for all variables in a model.

Example: For variables like "Weather" (Sunny, Rainy), "Cold" (Yes, No), and "Toothache" (Yes, No), the full joint probability distribution shows the probability for every combination, like "Sunny, Cold = Yes, Toothache = No."

Key Terms

Continuous Probability Distributions

For continuous variables like temperature, we can't say the probability of it being exactly a specific value (like 20.5°C) because there are infinitely many possible values. Instead, we use a density function to describe how probabilities are spread across a range of values.

Probability Density Functions: For continuous variables, we use density functions. These don't give the probability of a specific value but describe the likelihood within a range.

Example of Continuous Probability
If tomorrow's temperature is between 18°C and 26°C, a density function might say that the temperature is evenly distributed between these values. If the density function is 0.125, it means that within a small range around a certain temperature, the likelihood of falling into that range is related to this density value.

Conditional Probability

Ms.Rafath Zahra

Conditional probability is about figuring out how likely something is given that we know some other information.

Example: Imagine you know a patient has a toothache. The conditional probability tells us how likely it is that the patient also has a cavity, based on this information.

Notation: Written as P(a∣b), it means “the probability of a given that b is true.”

Posterior Probability: This is the updated probability of an event happening after considering new evidence.

Example: If you find out a patient has a toothache, the updated chance (posterior probability) of them having a cavity might be higher than if you didn’t know about the toothache.

Product Rule : This rule helps calculate the probability of two events happening together by breaking it down into simpler parts.

Example: To find the probability of both having a cavity and a toothache, you can use the product rule:

$P(a \text{ and } b) = P(a|b) \times P(b)$

This means you first find the probability of having a cavity given a toothache and then multiply it by the probability of having a toothache.

Be careful not to confuse conditional probabilities with logical certainty. Just because $P(a|b)=0.8$ (the probability of having a cavity given a toothache) doesn’t mean that every time someone has a toothache

Probability Axioms: Kolmogorov's Axioms

Ms.Rafath Zahra

1. Probability Range:

- The chance of any event happening is between 0 and 1.
- 0 means the event will definitely not happen.
- 1 means the event will definitely happen.
- Example: If you flip a fair coin, the probability of landing heads is 0.5 (or 50%).

2. Certain and Impossible Events:

- If something always happens, its probability is 1 (it's certain).
- If something never happens, its probability is 0 (it's impossible).
- Example: The probability of the sun rising tomorrow is very close to 1 because it's certain to happen.

3. Probability of Combined Events (Disjunction):

If you want to find the probability of either of two events happening (like A or B), you need to consider both events but avoid counting any overlap twice.

Formula:

$P(A \text{ or } B)=P(A)+P(B)-P(A \text{ and } B)$

Example: If the chance of rain today is 0.3 and the chance of snow is 0.2, but there's a chance both could happen together with a probability of 0.05, the combined chance of either rain or snow is

$0.3+0.2-0.05=0.45$.

Bayes' Rule

Ms.Rafath Zahra

Bayes' Rule helps us update our beliefs about the probability of an event based on new evidence. It allows us to use known probabilities to calculate a more accurate probability when we get new information.

How Does It Work?
Imagine you have two things you're interested in:

A disease (like meningitis)
A symptom (like a stiff neck)
Bayes' Rule helps answer questions like, "If a patient has a stiff neck, what’s the probability they have meningitis?"

The Formula
The formula for Bayes' Rule looks like this:

$$P(\text{Disease} \mid \text{Symptom}) = \frac{P(\text{Symptom} \mid \text{Disease}) \times P(\text{Disease})}{P(\text{Symptom})}$$

Bayes' Rule

Ms.Rafath Zahra

Here's what each part means:

P(Disease | Symptom): The probability of having the disease given that you have the symptom.
P(Symptom | Disease): The probability of having the symptom if you have the disease.
P(Disease): The overall probability of having the disease.
P(Symptom): The overall probability of having the symptom.

Example: Known Information:

50% of people with meningitis have a stiff neck (i.e., P(Stiff Neck | Meningitis)=0.5).
The chance of any random person having meningitis is 1 in 50,000
(i.e., P(Meningitis)= 1 / 50,000
The chance of any random person having a stiff neck is 1 in 20 (i.e., P(Stiff Neck)= 1/20

Using Bayes' Rule:

$$P(\text{Meningitis} \mid \text{Stiff Neck}) = \frac{P(\text{Stiff Neck} \mid \text{Meningitis}) \times P(\text{Meningitis})}{P(\text{Stiff Neck})}$$

$$P(\text{Meningitis} \mid \text{Stiff Neck}) = \frac{0.5 \times \frac{1}{50{,}000}}{\frac{1}{20}}$$

$$P(\text{Meningitis} \mid \text{Stiff Neck}) \approx \frac{0.5 \times 0.00002}{0.05} = 0.0002 \text{ or } 0.02\%$$

Dempster-Shafer theory

Ms.Rafath Zahra

Belief Function: Instead of just calculating the probability of something happening, Dempster-Shafer theory calculates how strongly the evidence supports that thing. This measure is called a "belief function."

Example with Coins:

Initial Ignorance: If you don’t know if a coin is fair or not, your belief about heads or tails is uncertain. Initially, you might say you have no strong belief either way.
Expert Input: If an expert is 90% sure the coin is fair (meaning it has a 50% chance of heads), you use this to update your belief. For instance, your belief that the coin will land heads could be 0.45 based on the expert's input.

Decision Making:
Conflicting Evidence: If your belief function says you don't have strong evidence for or against the coin coming up heads, it may be hard to make decisions based on it.
Probabilistic Comparison: Unlike probability, where you make decisions based on expected outcomes, Dempster-Shafer theory may not always lead to clear decisions, especially when there's conflicting or insufficient evidence.

Probability Interval:

Range of Belief: Dempster-Shafer theory often shows a range of beliefs (e.g., [0,1] before expert input and [0.45, 0.55] after). This range indicates the uncertainty or how much more evidence you might need.

Ms.Rafath Zahra

THANK YOU

MEASI
Institute of
Information Technology