Module 5

Uncertain Knowledge and Reasoning:
Quantifying Uncertainty:

Contents

1. Acting Under Uncertainty

Agents Under Uncertainty

- **Partial Observability :** the agent can't observe everything
- **Nondeterminism**: the environment's behavior is not completely predictable
- **Agents** often encounter uncertainty due to partial observability or nondeterminism.
- **Strategies** like maintaining a **belief state** or **generating contingency plans** help handle this uncertainty.

Belief State

- Is essentially a **probability distribution** over possible states of the environment.
- This distribution represents the agent's beliefs about the likelihood of being in each possible state, given the ***available evidence and observations***.
- By maintaining a belief state, the agent can make decisions that are **robust to uncertainty**. It can update its **belief state** as it receives **new information**, allowing it to adapt its strategies and actions accordingly.

Contingency plans

- A contingency plan is **a predefined course of action** designed for agents to take actions depending on how the environment evolves.
- These plans provide **flexibility and resilience** in dealing with uncertainty, allowing the agent to adapt to different scenarios as they unfold.

Limitations

1. **Logical agents** must consider all possible explanations for observations, leading to **complex belief-state** representations.
2. **Contingency plans** can become excessively large and must account for improbable events.
3. **Sometimes, there's no guaranteed plan** for achieving a goal, yet the agent must act.

Example

- For instance, an **automated taxi** aiming to get a passenger to the airport on **time may face uncertainty** regarding various factors like **traffic or accidents**.
- Despite uncertainty, the agent must make decisions that maximize its expected **performance, considering its knowledge about the environment**. The right decision depends on balancing goals and their likelihood of achievement.
- This highlights the importance of **rational decision-making** under uncertainty, where agents choose actions that are expected to optimize their performance measures relative to their knowledge about the environment.

Summarizing Uncertainty

- Consider the scenario of diagnosing a dental patient's toothache, which exemplifies uncertain reasoning
- **Rule1: "Toothache ⇒ Cavity"** : is inadequate because not all patients with toothaches have cavities; some may have **gum disease**, **abscesses**, or other issues.
- To rectify this, one would need to enumerate an extensive list of possible problems, making the rule impractical
- **Rule2 : Toothache ⇒ Cavity ∨ GumProblem ∨ Abscess ...**

Summarizing Uncertainty

- Even transforming the rule into a causal one like:
- **"Cavity ⇒ Toothache"**
- proves inaccurate because not all **cavities** cause pain.

The limitations of using logic for medical diagnosis

1. **Laziness**: It's cumbersome to list all antecedents or consequents required for a rule, and using such exhaustive rules is challenging.
2. **Theoretical ignorance**: Medical science lacks a complete theory for the domain.
3. **Practical ignorance:** Even with knowledge of rules, uncertainty remains due to incomplete testing or unavailable tests.

2. Basic Probability Notation

Use of Probability

- Probability allows summarizing uncertainty, indicating, for example, an **80% chance that a patient** with a **toothache has a cavity**. This belief could derive from **statistical data or general dental knowledge**.
- **Probability statements** reflect knowledge states rather than absolute truths about the real world.
- For instance, stating "**The probability of a cavity, given a toothache, is 0.8**" reflects a **certain knowledge state**.
- As **new information emerges**, such as a history of gum disease, the probability statement may change accordingly without contradicting previous assessments.

Uncertainty and rational decisions

- Preferences, expressed through utilities, are combined with probabilities in the general theory of rational decisions called decision theory:
- **Decision theory=probability theory + utility theory**
- The fundamental idea of decision theory is that an **agent i**s rational if and only if it chooses the action that yields the **highest expected utility,** averaged over all the **possible outcomes of the action**. This is known as the **principle of maximum expected utility (MEU).**

```
function DT-AGENT(percept) returns an action
    persistent: belief_state, probabilistic beliefs about the current state of the world
                action, the agent's action

    update belief_state based on action and percept
    calculate outcome probabilities for actions,
        given action descriptions and current belief_state
    select action with highest expected utility
        given probabilities of outcomes and utility information
    return action
```

Function Name: DT-AGENT (percept)

Purpose: To make decisions in an uncertain environment based on observed precepts.

Persistent Variables:

- **belief state**: Probabilistic beliefs regarding the current state of the environment.
- **action**: Represents the action chosen by the agent.

Functionality:

- Updates the belief state based on the agent's action and the received percept.
- Calculates the probabilities of different outcomes for each potential action, considering action descriptions and the current belief state.
- Selects the action with the highest expected utility, taking into account the probabilities of outcomes and utility information associated with each outcome.

- **Returns the selected action**.

What probabilities are about?

- **Probability** is a **quantitative measure of the likelihood or chance** that a particular event or outcome will occur within a given set of circumstances. It is expressed as a numerical value between 0 and 1, where 0 indicates impossibility (the event will not occur), 1 indicates certainty (the event will occur), and values between 0 and 1 represent varying degrees of likelihood.
- In more formal terms, probability is defined as ***the ratio of the number of favorable outcomes to the total number of possible outcomes in a given event space.***
    - **P(Favorable_Outcomes)= No of Favorable_Outcomes/Total possible Outcomes**
- It provides a way to ***quantify uncertainty and make informed decisions*** based on the likelihood of different outcomes.
- Probability theory is widely applied in various fields, including *mathematics, statistics, science, engineering, economics, and social sciences, to analyze and predict the occurrence of events* and to ***make decisions under uncertainty.***

What probabilities are about?

- Probabilistic assertions are about **possible worlds**
- Probabilistic assertions talk about how **probable** the various worlds are?
- **Sample Space** : In probability theory, the set of all possible worlds is called the sample space.

Probability Model

- **Probability Model:** A fully specified probability model associates a numerical probability P(ω) with each possible world.
- The basic axioms of probability theory say that every possible world has a probability between 0 and 1 and that the total probability of the set of possible worlds is 1:

$$0 \leq P(\omega) \leq 1 \text{ for every } \omega \text{ and } \sum_{\omega \in \Omega} P(\omega) = 1 \, .$$

Events:

- Events in probability theory are **sets of possible worlds** that satisfy **certain conditions or criteria**. Events are typically described using propositions in a formal language.
- The probability associated with an **event** is defined as the sum of the probabilities of the worlds in which it holds.
- For example, **the event of rolling doubles** when rolling fair dice consists of the worlds (1,1), (2,2), ..., (6,6), and its probability is the sum of the probabilities of these individual worlds.

Unconditional/Prior Probability:

- Probabilities such as **P(Total = 11) and P(doubles)** are called unconditional or prior probabilities (and sometimes just “priors” for short);
- They refer to **degrees of belief** in propositions in the absence of any other information.
- Most of the time, however, we have some information, usually called **evidence,** that has already been revealed

Example

- Suppose you are a teacher and you want to estimate the probability that a **randomly selected student** in your class will **score above 80%** on the upcoming math test. **Before any students have taken the test**, you might use your knowledge of the class, previous test scores, and other relevant factors to estimate this probability.
- Let's say, based on your experience and knowledge of the class's performance, you believe that **about 30% of the students are likely to score above 80% on the math test.**
- In this example:
  - **Event A**: A student scores above 80% on the math test.
  - **Prior Probability $P(A)$** = 0.30 (or 30%)

Conditional Probability/Posterior Probability

- Conditional probability refers to the probability of an event occurring given that another event has already occurred or is known to have occurred. It is denoted by expressions like **P(A|B), where A is the event of interest and B represents the condition.**
- **Example:**
  - **P(doubles | Die1 = 5),**
  - **P(cavity |toothache) = 0.6**

Mathematically speaking, **conditional probabilities are defined in terms** of unconditional probabilities as follows: for any propositions **a** and **b**, we have:

$$P(a \mid b) = \frac{P(a \wedge b)}{P(b)}\,, \tag{13.3}$$

which holds whenever $P(b) > 0$. For example,

$$P(doubles \mid Die_1 = 5) = \frac{P(doubles \wedge Die_1 = 5)}{P(Die_1 = 5)}\,.$$

The definition of conditional probability, Equation (13.3), can be written in a different form called the **product rule**:

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

The product rule is perhaps easier to remember: it comes from the fact that, for a and b to be true, we need b to be true, and we also need a to be true given b.

The language of propositions in probability assertions

- Variables in probability theory are called random variables and their names begin with an uppercase letter
  - Total and Die1 are random variables
- Every random variable has a **domain**—the set of possible values it can take on.
  - The domain of Total for two dice is the set {2, . . . , 12} and
  - The domain of Die1 is {1, . . . , 6}
- A Boolean random variable has the domain **{true, false}**
- Domains can be sets of arbitrary tokens;
  - The domain of Age may be *{juvenile, teen, adult}* and
  - The domain of Weather might be *{sunny, rain, cloudy, snow}*.

The language of propositions in probability assertions

- Variables can have **infinite domains**, too—either discrete (like the integers) or continuous (like the reals).
- For any variable with an **ordered domain**, inequalities are also allowed, such as **NumberOfAtomsInUniverse ≥ 10^70** .
- We can combine **elementary propositions** by using the connectives of propositional logic.
- For example, we can express "**The probability that the patient has a cavity, given that she is a teenager with no toothache, is 0.1**" as follows:
  - **P(cavity | ¬toothache ∧ teen) = 0.1 .**

The language of propositions in probability assertions

- We could write:
  - P(Weather = sunny) = 0.6
  - P(Weather = rain) = 0.1
  - P(Weather = cloudy) = 0.29
  - P(Weather = snow) = 0.01 ,
- As P(Weather) = <0.6, 0.1, 0.29, 0.01>
- We say that the P statement defines a probability distribution for the random variable Weather .

The language of propositions in probability assertions

- The P notation is also used for conditional distributions:
  - P(X | Y) gives the values of P(X = xi | Y = yj) for each possible i, j pair
- For continuous variables, it is not possible to write out the entire distribution as a vector, because there are infinitely many values. Instead, we can define the probability that a random variable takes on some value x as a parameterized function of x.
- For example, the sentence given below expresses the belief that the temperature at noon is distributed uniformly between 18 and 26 degrees Celsius. **We call this a probability density function**.

$$P(NoonTemp = x) = Uniform_{[18C,26C]}(x)$$

The language of propositions in probability assertions

- Probability density functions (sometimes called pdfs) differ in meaning from discrete distributions. Saying that the probability density is uniform from 18C to 26C means that there is a 100% chance that the temperature will fall somewhere in that 8C-wide region and a 50% chance that it will fall in any 4C-wide region, and so on

$$P(NoonTemp = x) = Uniform_{[18C,26C]}(x)$$

The language of propositions in probability assertions

- We write the probability density for a continuous random variable X at value x as **P(X = x) or just P(x);** the intuitive definition of P(x) is the probability that X falls within an arbitrarily small region beginning at x, divided by the width of the region:

$$P(x) = \lim_{dx \to 0} P(x \leq X \leq x + dx)/dx \ .$$

For *NoonTemp* we have

$$P(NoonTemp = x) = Uniform_{[18C,26C]}(x) = \begin{cases} \frac{1}{8C} \text{ if } 18C \leq x \leq 26C \\ 0 \text{ otherwise} \end{cases} ,$$

The language of propositions in probability assertions

- ***Joint Probability Distributions :*** In addition to distributions on single variables, we need notation for distributions on multiple variables. Commas are used for this.
- For example**, P(Weather , Cavity)** denotes the probabilities of all combinations of the values of **Weather and Cavity**.
- This is a 4 × 2 table of probabilities called the joint probability distribution of Weather and Cavity.
- We can also mix variables with and without values; **P(sunny, Cavity)** would be a **two-element vector** giving the probabilities of a **sunny day with a cavity and a sunny day with no cavity**

The language of propositions in probability assertions

- The **P notation** makes certain expressions much more concise than they might otherwise be.
- For example, the product rules for all possible values of Weather and Cavity can be written as a single equation:
  - **P(Weather , Cavity) = P(Weather | Cavity)P(Cavity)**
  - instead of as these 4 × 2 = 8 equations (using abbreviations W and C):

$$
\begin{aligned}
P(W = sunny \wedge C = true) &= P(W = sunny | C = true)\, P(C = true)\\
P(W = rain \wedge C = true) &= P(W = rain | C = true)\, P(C = true)\\
P(W = cloudy \wedge C = true) &= P(W = cloudy | C = true)\, P(C = true)\\
P(W = snow \wedge C = true) &= P(W = snow | C = true)\, P(C = true)\\
P(W = sunny \wedge C = false) &= P(W = sunny | C = false)\, P(C = false)\\
P(W = rain \wedge C = false) &= P(W = rain | C = false)\, P(C = false)\\
P(W = cloudy \wedge C = false) &= P(W = cloudy | C = false)\, P(C = false)\\
P(W = snow \wedge C = false) &= P(W = snow | C = false)\, P(C = false)\ .
\end{aligned}
$$