Probabilistic Reasoning

Bridging Knowledge-Based Systems and Real-World Data

Rule-Based Systems Vs. Real World

Expert systems and Ontologies

- **Rule-based**
- **Structured approaches**

Search Algorithms (Classical & Adversarial)

- **Rely on deterministic or game-based logic.**

What about Real-world AI ?

Rule-Based Systems Vs. Real World

Complex, Noisy Data

- Real-world environments (sensors, user inputs, big data) introduce noise or errors.
- **Example**: A self-driving car’s sensors can be obstructed by weather, leading to uncertain measurements.

Incomplete Information

- Agents rarely have full knowledge of the world state.
- **Example:** A medical diagnosis system does not always have results for every possible test.

Rule-Based Systems Vs. Real World

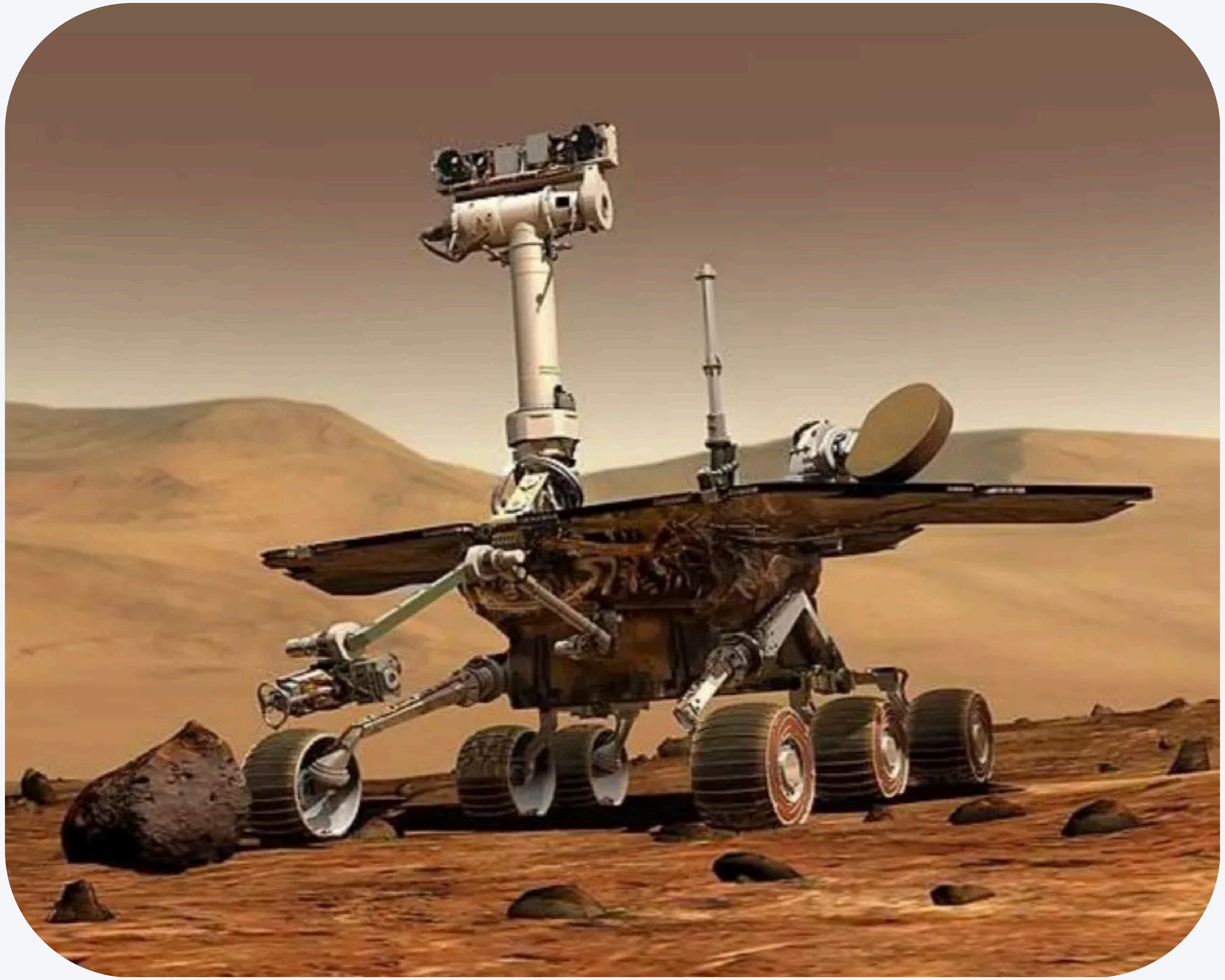
Uncertain Outcomes

- **Decisions** often need to be made with partial data (e.g., is this patient likely to have a disease?).

Dynamic & Changing Environments

- The world is not static: conditions and user behaviors evolve over time
- Think e-commerce trends or climate changes

Rule-Based Systems Vs. Real World

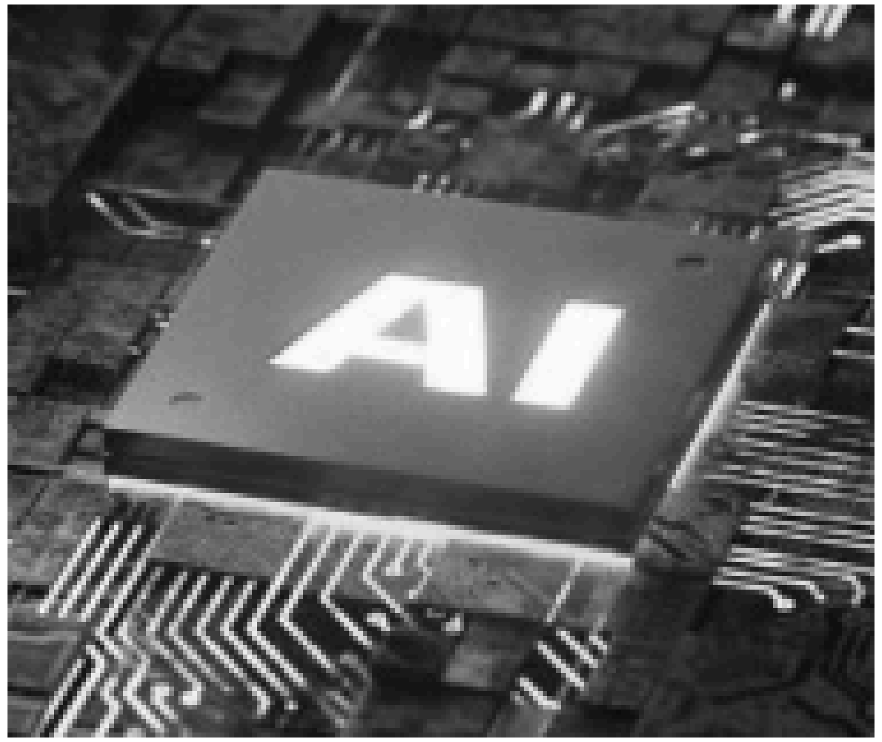


Rule-Based Systems Vs. Real World

Rule-Based Systems Vs. Real World

Performance uncertainty

How will autonomus vehicles perform in sunny and snow wealther?



Technology uncertainty

When will AI applications be smarter than human controlled ones?

Implication uncertainty

Will AI really render all jobs obsolete?

AI’s Shift from Rules to Data

AI’s Shift from Rules to Data

Real-world AI often encounters **imperfect or incomplete information**

Foundations for Modern AI:

- Almost all modern AI applications (computer vision, NLP, recommender systems, etc.) rely on machine learning to make predictions or decisions under **uncertainty**.
- **Machine Learning** is the dominant approach—algorithms learn patterns from data rather than relying solely on hand-coded rules.

AI’s Shift from Rules to Data

Uncertainty as a Core Challenge:

- Almost all modern AI fields—like computer vision, natural language processing, and recommender systems—must deal with uncertain or ambiguous inputs.

Examples of AI Under Uncertainty:

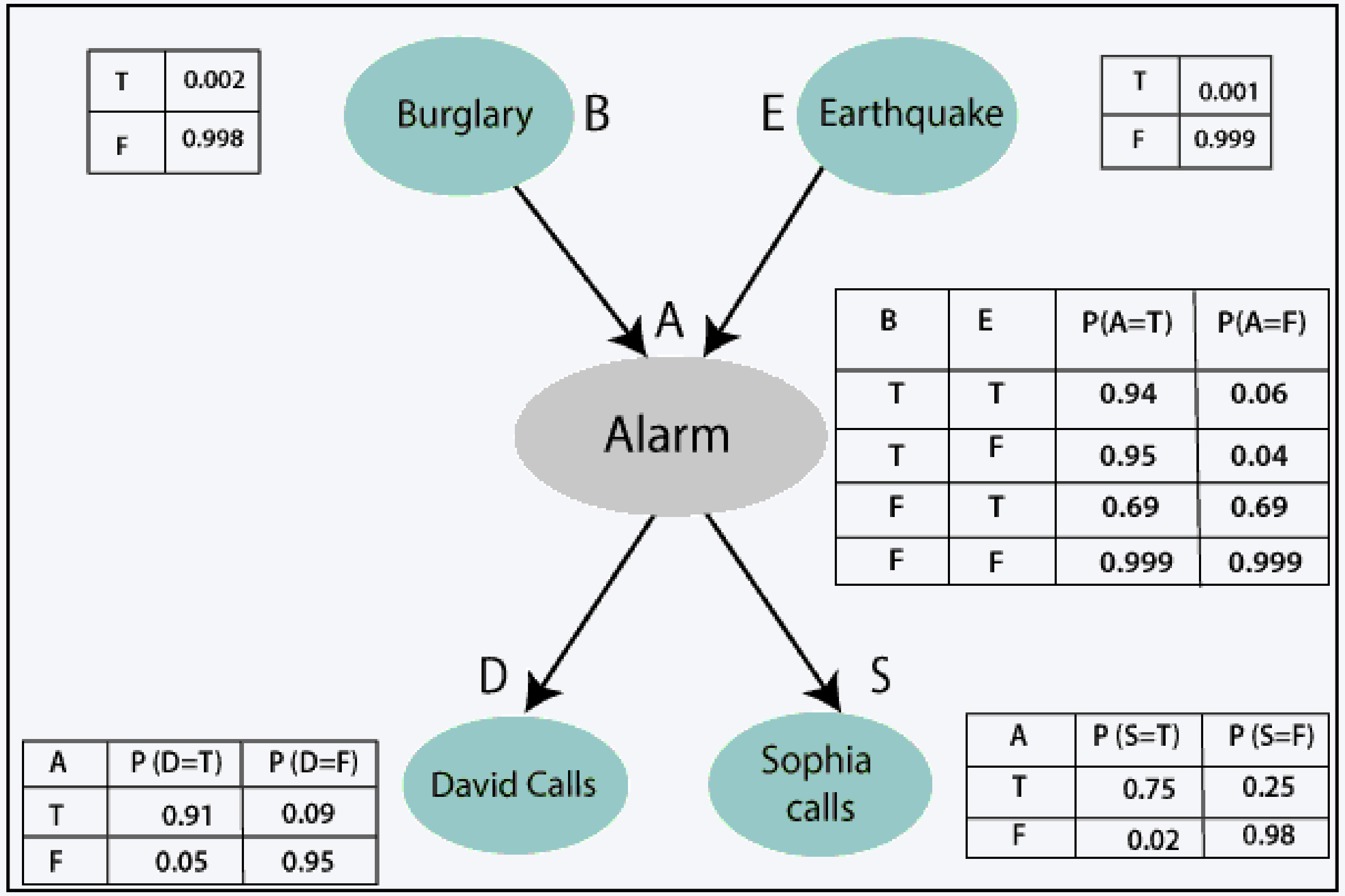
- **Computer Vision:** A system classifying images in low light or occluded scenes must account for noise and partial views.
- **Natural Language Processing:** Speech recognition deals with background noise and variations in accents.
- **Recommender Systems:** Predictions rely on partial user data, uncertain user preferences, and shifting trends.

From Data to Decisions

From Data to Decisions

- **Machine learning** uses statistical methods (including Bayesian techniques) to cope with incomplete or imperfect data.

- **Probabilistic models** (e.g., Naive Bayes, Hidden Markov Models, Bayesian Networks) are used for classification, time-series prediction, and more.

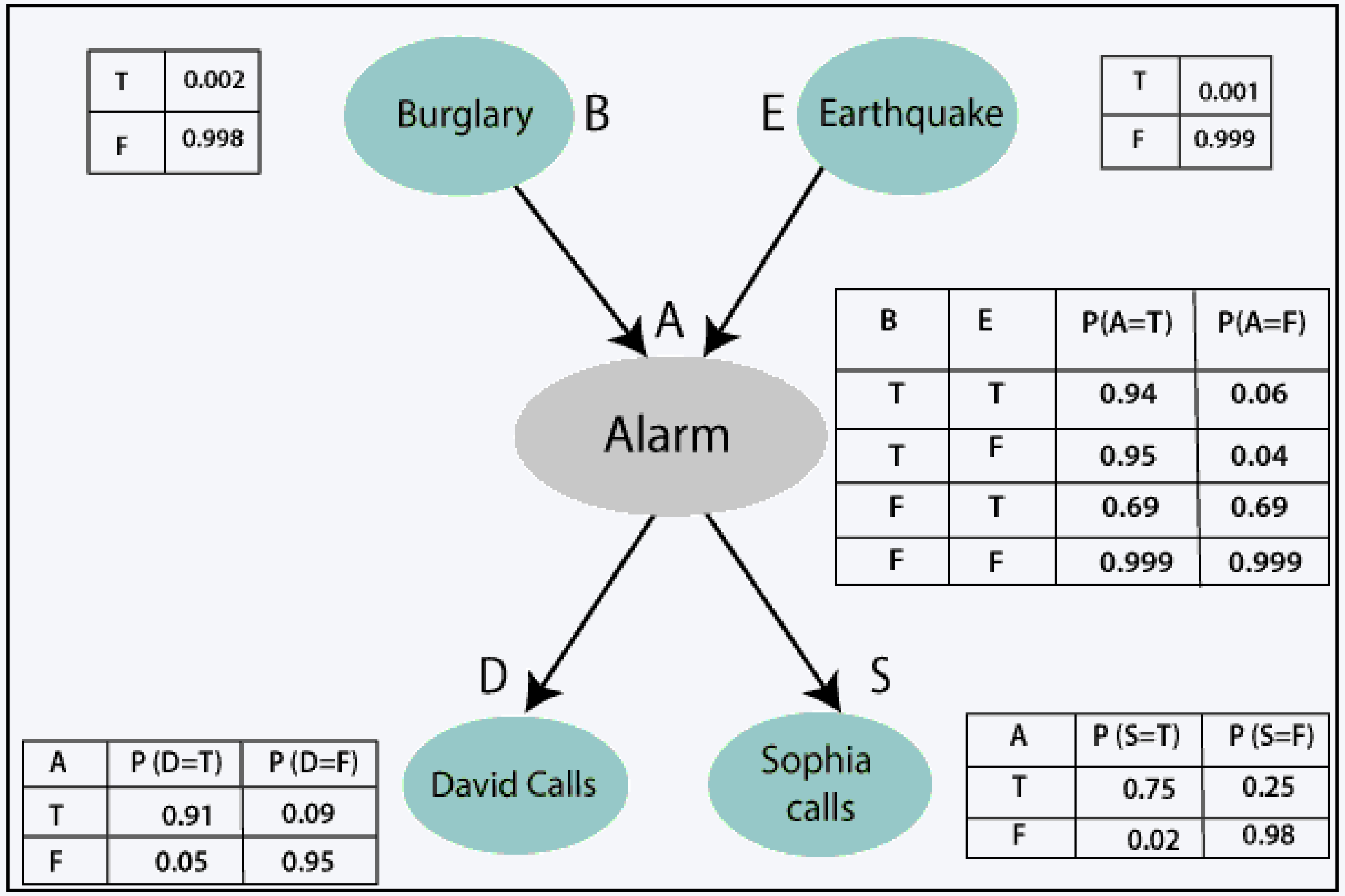


Example on Konwledge representation using Bayesian networks

Bayesian Belief Network in Artificial Intelligence
Bayesian Belief Network in Artificial Intelligence with Tutorial, Introduction, History of Artificial Intelligence, AI, AI Overview, Application of AI, Types of AI, What is AI, subsets of ai, types of agents, intelligent...
javatpoint.com

Why Probabilistic Models?

- **Inherent Uncertainty:** Sensors, user inputs, or natural phenomena (like weather) can produce ambiguous data.
- **Adaptive Learning:** Probabilistic models can update their estimates whenever new evidence appears, making them more robust to fluctuations in data quality.
- **Decision-Making Under Risk:** These methods compute the probability of various outcomes and let a system select the action that maximizes some measure of success, despite uncertainty.



Example on Konwledge representation using Bayesian networks

Bayesian Belief Network in Artificial Intelligence

Bayesian Belief Network in Artificial Intelligence with Tutorial, Introduction, History of Artificial Intelligence, AI, AI Overview, Application of AI, Types of AI, What is AI, subsets of ai, types of agents, intelligent...

javatpoint.com

Why This Matters?

- **Robustness:** Even with noisy, incomplete, or evolving data (like constantly changing spam strategies), the Naive Bayes approach remains effective by accumulating evidence across many features.

- **Scalability:** Probabilistic methods can be extended or combined with other techniques (e.g., advanced NLP tools) as data grows.

- **Continuous Learning:** Because these methods are rooted in probability, they offer a natural mechanism for updating beliefs when the environment or data distribution shifts.

AI Uncertainty

Uncertainty

- Suppose A and B are two statements, If we implement **if-then** rule to these statements, we might write **A→B**, Which means if **A is true** Then **B is true**, or if **A is false** then **B is false.**

- **But** consider a situation where we are not sure about **whether A is true or false >> So** we cannot express this statement, this situation is called **uncertainty**.

- So to represent **uncertain knowledge**, where we are not sure about the **predicates**, **>>>** we need **uncertain reasoning** or **probabilistic reasoning.**

Causes of uncertainty

1. Information occurred from unreliable sources.

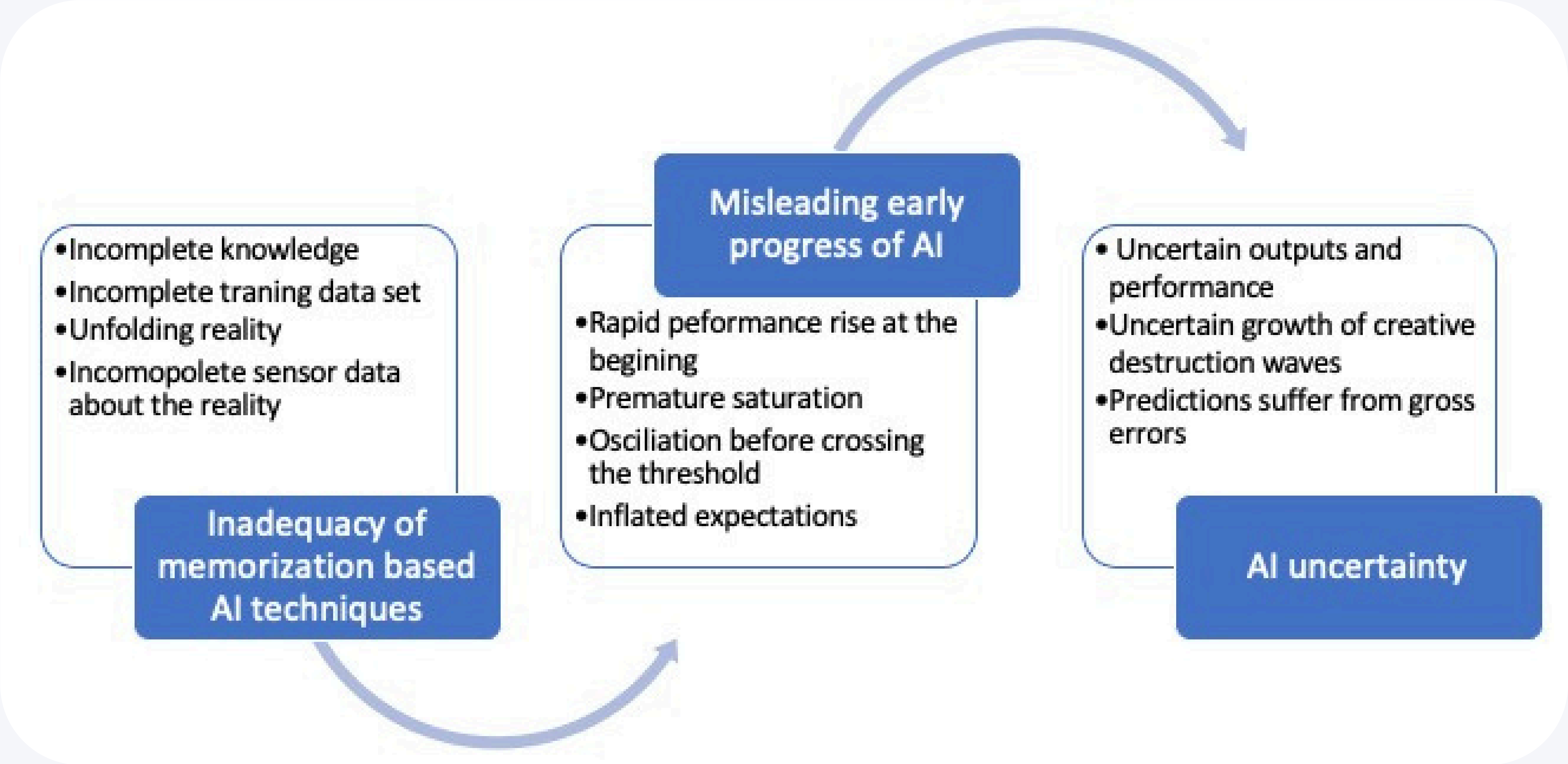
2. Experimental Errors

3. Equipment fault

4. Temperature variation

5. Climate change.

Causes of uncertainty



https://www.the-waves.org/2023/11/06/ai-uncertainty-from-technology-to-liability/

Probabilistic Reasoning

Probabilistic Reasoning

- **Probabilistic reasoning** is a way of **knowledge representation** where we apply the concept of **probability** to indicate the **uncertainty in knowledge.**

- In probabilistic reasoning, we combine **probability theory with logic** to handle the **uncertainty**.

Need of probabilistic reasoning in AI

  - When there are unpredictable outcomes.
  - When specifications or possibilities of predicates becomes too large to handle.
  - When an unknown error occurs during an experiment.

Probabilistic Reasoning

In probabilistic reasoning, there are two ways to solve problems with uncertain knowledge:

- **Bayes' rule**
- **Bayesian Statistics**

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}}$$

Bayes' rule

Probability

A

$$P(A)$$

where P(A) is the probability of an event A.

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

$$P(A) = 0$$

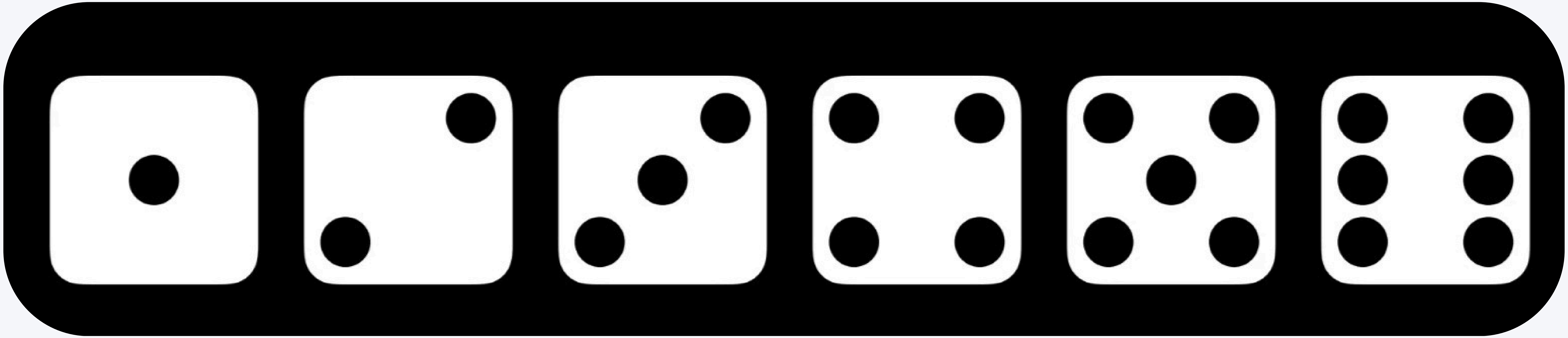
indicates total uncertainty in an event A.

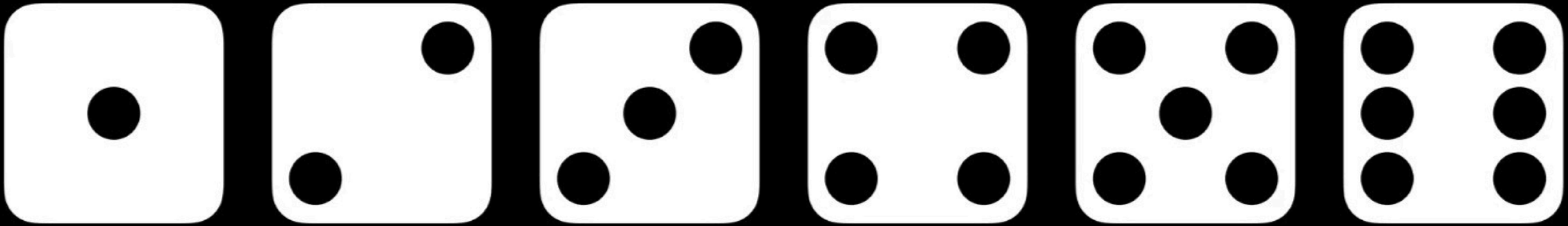
$$P(A) = 1$$

indicates total certainty in an event A.

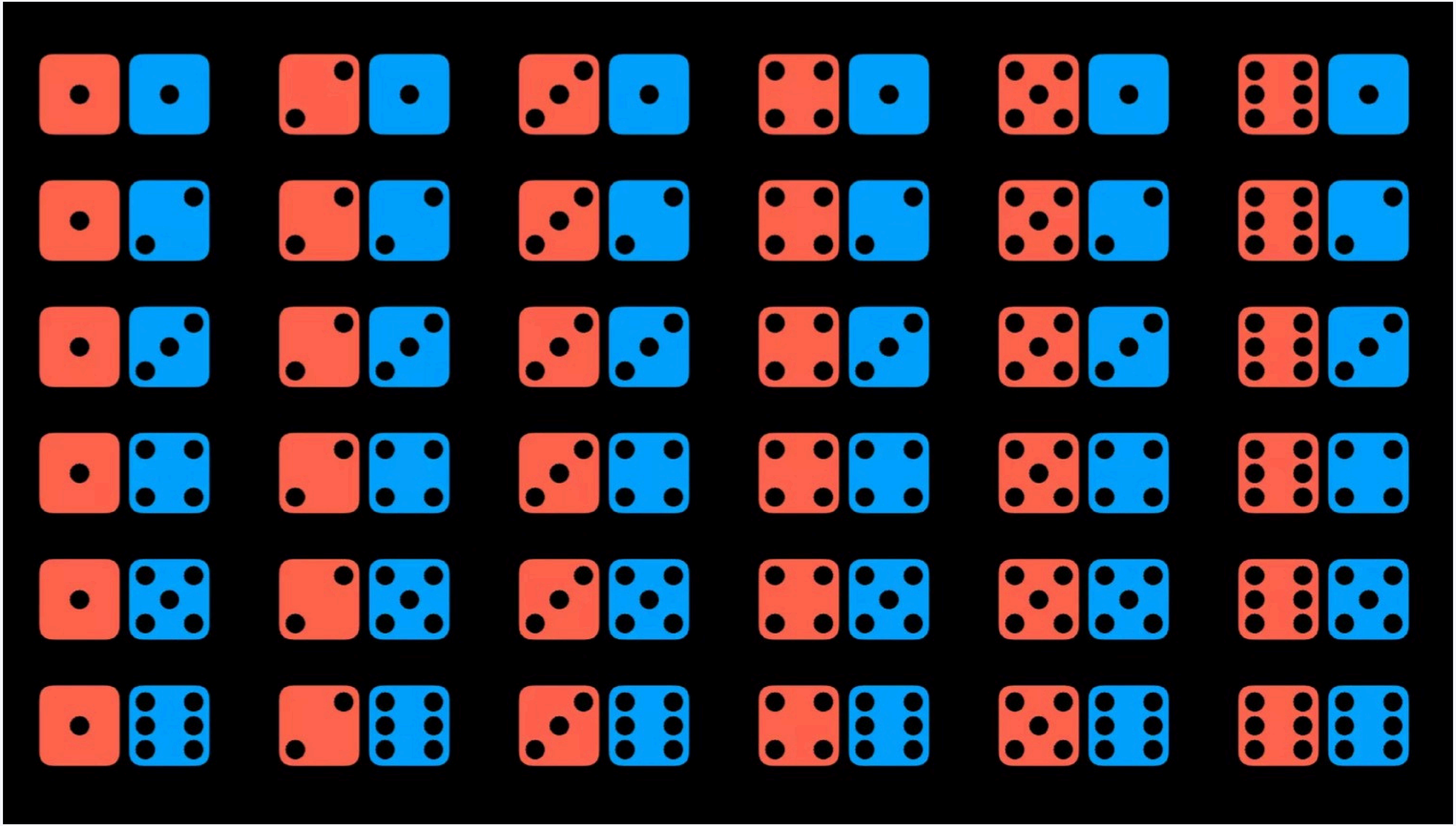
$$\sum P(A_i) = 1$$

Sum of probabilities





$\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$



2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12


2 3 4 5 6 7
3 4 5 6 7 8
4 5 6 7 8 9
5 6 7 8 9 10
6 7 8 9 10 11
7 8 9 10 11 12

$$P(\text{Sum} = 7) =$$

$$P(\text{Sum} = 12) =$$

$$P(Sum = 7) = 6/36 >> 1/6$$

$$P(Sum = 12) = 1/36$$

Probability

Probability can be defined as a chance that an uncertain event will occur. It is the numerical measure of the likelihood that an event will occur.

The value of **probability** always remains between **0 and 1** that represent ideal uncertainties.

1. **$0 \le P(A) \le 1$,** where P(A) is the probability of an event A.
2. **$P(A) = 0$,** indicates total uncertainty in an event A.
3. **$P(A) = 1$,** indicates total certainty in an event A.

Probability

We can find the **probability** of an uncertain event by using the below formula.

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

- **P(¬A) = probability of a not happening event.**
- **P(¬A) + P(A) = 1.**

Probabilistic Reasoning

- **Event**: Each possible outcome of a variable is called an event. **Sample space:** The collection of all possible events is called sample space. **Random variables:** Random variables are used to represent the events and objects in the real world. **Prior probability:** The prior probability of an event is probability computed before observing new information. **Posterior Probability:** The probability that is calculated after all evidence or information has taken into account. It is a combination of prior probability and new information.

Conditional probability

- **Conditional probability** is a probability of occurring an event when another event has already happened.

Let's suppose, we want to calculate the **event A when event B has already occurred**, Then, **“the probability of A under the conditions of B"**, it can be written as:

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$

Where **P(A⋀B)**= Joint probability of **A and B**
P(B)= Marginal probability of **B.**

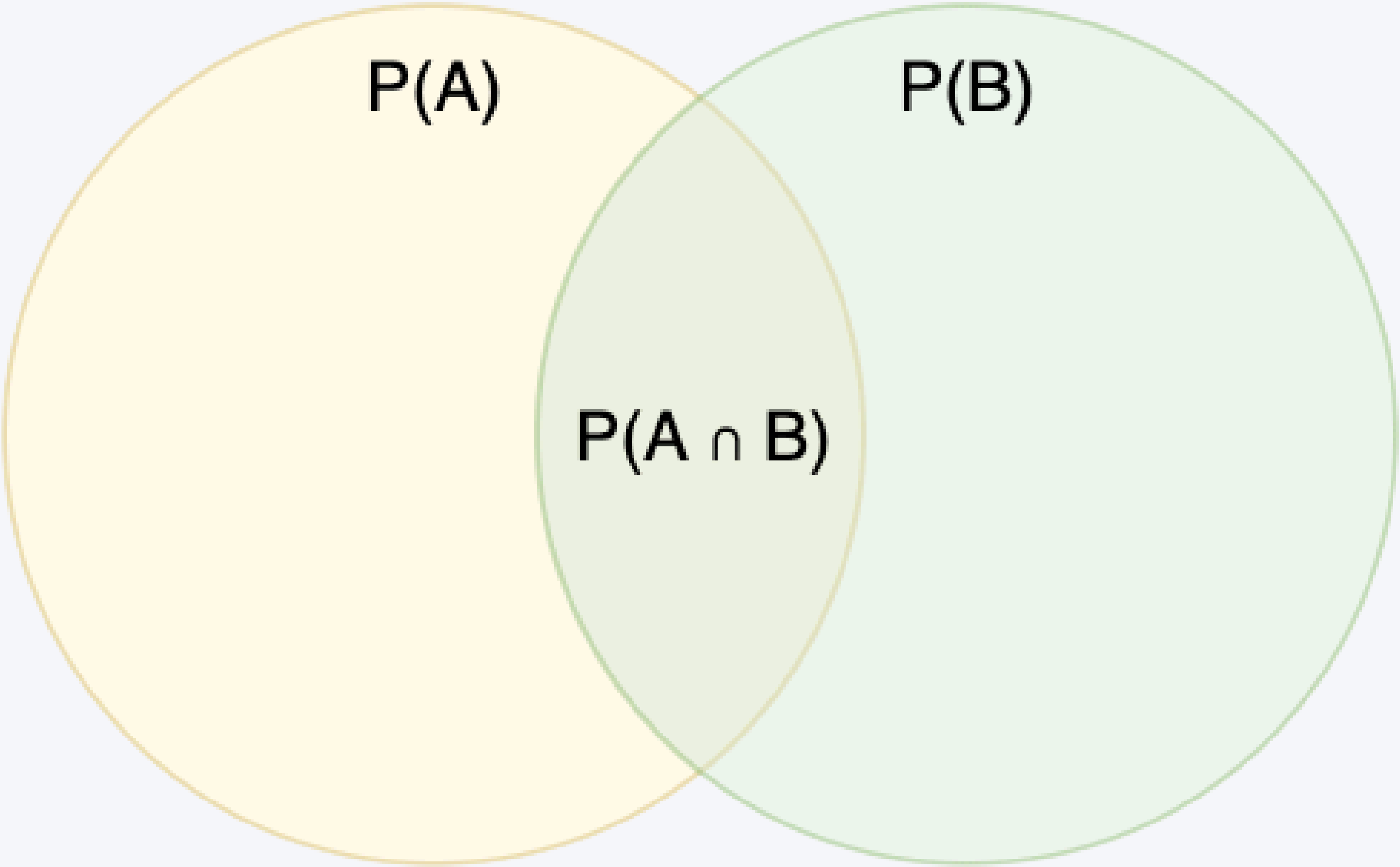
Conditional probability

P(rain today | rain yesterday)

P(route change | traffic conditions)

P(disease | test results)

Venn Diagram



Venn Diagram Example

In a class, there are **70% of the students** who **like English** and **40% of the students** who **likes English and mathematics**, and

.... Then what is the percent of students those who like English also like mathematics?

in other words,

What is the probability of (Students like Math if they like English)

Venn Diagram Example

In a class, there are **70% of the students** who **like English** and **40% of the students** who **likes English and mathematics**, and

.... Then what is the percent of students those who like English also like mathematics?

- Let, A is an event that a student likes Mathematics
- B is an event that a student likes English

Venn Diagram Example

$$P(A|B) = \frac{P(A \land B)}{P(B)} = \frac{0.4}{0.7} = 57\%$$

- Hence, 57% are the students who like English also like Mathematics.

Bayes' theorem

- Bayes' Theorem is a mathematical rule that helps us figure out the probability of something happening based on new information, given what we already know.
  - Is it going to **rain today**, I can see clouds in the sky.
  You know how often it rains on any given day in your city. Let's say it rains **20%** of the time **P(A)**
  -

Incorporate new evidence:

  - You look up and see **clouds**. You know that when it rains, it's **cloudy 90%** of the time **P(B | A)**.
  - But you also know that it's **cloudy even when it doesn't rain**, say **50%** of the time **P(B | notA))**.
  Combine the two: **Bayes' Theorem** combines this information to update the probability of rain today, taking the clouds into account.
  -

Bayes' theorem

Bayes' Theorem is a mathematical rule that helps us figure out the probability of something happening based on new information, given what we already know.

Here’s the formula:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Let’s break it down:

- P(A ∣ B): The probability of event AA happening, given that event BB has already happened. This is what we are trying to find.
- P(B ∣ A): The probability of event BB happening if AA is true.
- P(A): The overall probability of AA happening (before considering BB).
- P(B): The overall probability of BB happening.

Bayes' theorem

Bayes' theorem can be derived using product rule and conditional probability of event A with known event B:

As from product rule we can write:

$P(A \wedge B) = P(A|B)\,P(B)$

or

Similarly, the probability of event B with known event A:

$P(A \wedge B) = P(B|A)\,P(A)$

Equating right hand side of both the equations, we will get:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayes' theorem

- It shows the simple relationship between joint and conditional probabilities. Here,
- **P(A|B) is known as posterior**, which we need to calculate, and it will be read as Probability of hypothesis A when we have occurred an evidence B.

  P(B|A) is called the likelihood, in which we consider that hypothesis is true, then we calculate the
- probability of evidence.

  P(A) is called the prior probability, probability of hypothesis before considering the evidence

  P(B) is called marginal probability, pure probability of an evidence.
- In the equation (a), in general, we can write **P (B) = P(A)*P(B|Ai)**
-
-

$$P(A_i|B) = \frac{P(A_i)*P(B|A_i)}{\sum_{i=1}^{k} P(A_i)*P(B|A_i)}$$

Where A1, A2, A3,........, An is a set of mutually exclusive and exhaustive events

Application of Bayes' theorem in Artificial intelligence

Bayes' theorem in AI

Following are some applications of **Bayes' theorem:**

- It is used to calculate the **next step of the robot** when the already executed step is given.
- Bayes' theorem is helpful in **weather forecasting**.

Bayesian belief network is key computer technology for dealing with probabilistic events and to solve a problem which has uncertainty. We can define a Bayesian network as:

"A Bayesian network is a probabilistic graphical model which represents a set of variables and their conditional dependencies using a directed acyclic graph. "

- It is also called a **Bayes network, belief network, decision network, or Bayesian model.**
- Bayesian networks are **probabilistic**, because these networks are built from a probability distribution, and also use probability theory for prediction and anomaly detection.
- It can also be used in various tasks including prediction, anomaly detection, diagnostics, automated insight, reasoning, time series prediction, and decision making under uncertainty.

Bayes' theorem in AI

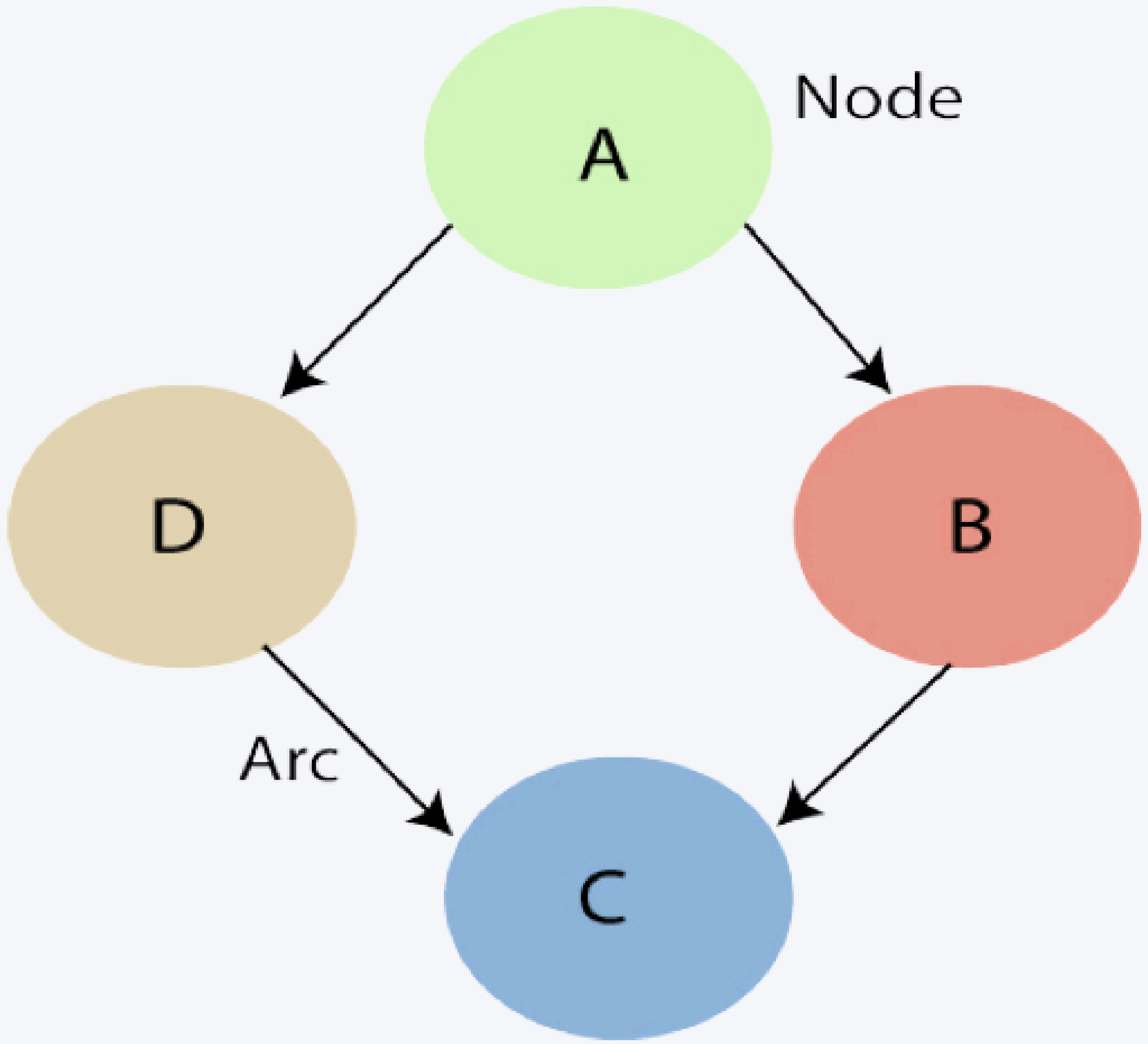
Bayesian Network can be used for building models from data and experts opinions, and it consists of two parts:

- **Directed Acyclic Graph**
- **Table of conditional probabilities.**

The generalized form of Bayesian network that represents and solve decision problems under uncertain knowledge is known as an Influence diagram.

A Bayesian network graph is made up of

- nodes
- Arcs (directed links)



Bayes' theorem in AI

Alarm **'A' – a node**, say installed in a house of a **person'X'**, which rings upon two probabilities i.e **burglary 'B' and Earthquake 'E'**, which are– parent nodes of the alarm node.

The alarm is the parent node of **two probabilities D calls 'D' & S calls 'S'** person nodes.

Upon the instance of burglary and Earthquake**, 'D' and 'S'** call person **'X'**, respectively.

But, there are few drawbacks in this case, as sometimes **'D' may forget to call the person'X',** even after hearing the alarm, as he has a tendency to forget things, quick.

Similarly, '**S', sometimes fails to call the person'X'**, as he is only able to hear the alarm, from a certain distance.