4. Inference and Reasoning

Artificial Intelligence and Neural Network (AINN)

Dr. Udaya Raj Dhungana
Assist. Professor
Pokhara University, Nepal
Guest Faculty
Hochschule Darmstadt University of Applied Sciences, Germany
E-mail: udaya@pu.edu.np and udayas.epost@gmail.com

Overview

- Reasoning- Deductive/Inductive
- Inference and rules of inference
- Forward and Backward Chaining
- Uncertainty
- Uncertain Reasoning
- Baye’s Theorem
- Bayesian Network
- Case-based Reasoning

Reasoning

- Mathematical reasoning involves
  - gathering evidence,
  - building arguments, and
  - drawing logical conclusions about these various ideas and their relationships.

Reasoning

- Reason is the capacity of consciously applying logic by drawing conclusions from new or existing information, with the aim of seeking the truth
- Reason is sometimes referred to as rationality.
- Reasoning is associated with the acts of thinking and cognition, and involves using one's intellect.
- The field of logic studies the ways in which humans can use formal reasoning to produce logically valid arguments
- Logic is the study of reasoning
- Types of reasoning:
    - deductive reasoning and
    - inductive reasoning.

Reasoning

Deductive	Inductive
Start with rules.	Establish a rule.
↓	↑
Make another rule.	Start with observations.

Reasoning

Deductive

Gravity makes things fall downwards.

Things that fall from a great height get hurt.

If I jump off this building, I will fall downwards.

Inductive

Things I throw off the roof fall down.

I threw a ball off the roof and it fell down.
I threw a rock off the roof and it fell down.
I threw a cat off the roof and it fell down.

Reasoning

- Deductive reasoning :
  - Deduction is a form of reasoning in which a conclusion follows necessarily from the stated premises.
  - A deduction is also the conclusion reached by a deductive reasoning process.
  - One classic example of deductive reasoning is that found in syllogisms like the following:

    Premise 1: All humans are mortal.
    Premise 2: Socrates is a human.
    Conclusion: Socrates is mortal.

  - The reasoning in this argument is deductively valid because there is no way in which the premises, 1 and 2, could be true and the conclusion, 3, be false.

Reasoning

- Inductive reasoning:
  - Induction is a form of inference producing propositions about unobserved objects or types, either specifically or generally, based on previous observation.
  - It is used to ascribe properties or relations to objects or types based on previous observations or experiences, or to formulate general statements or laws based on limited observations of recurring phenomenal patterns.
  - the truth of the premises does not guarantee the truth of the conclusion
  - Instead, the conclusion of an inductive argument follows with some degree of probability.

Reasoning

- Inductive reasoning:
  - Relatedly, the conclusion of an inductive argument contains more information than is already contained in the premises.
  - Thus, this method of reasoning is ampliative.
  - A classic example of inductive reasoning comes from the empiricist David Hume:

Premise: The sun has risen in the east every morning up until now.
Conclusion: The sun will also rise in the east tomorrow.

Reasoning

Types of Reasoning

DEDUCTIVE	VS	INDUCTIVE
States general idea and verifies to reach conclusion		Collects observations to reach general conclusion
Top down approach to problem solving		Bottoms up approach to problem solving
Example: Creating a social media marketing plan for Gen Z segment General Idea: performance results from the social media campaign will determine if social media is the best way to reach Gen Z'ers		Example: Using employee survey results to create a new employee wellness program Observations: feedback and results from employee survey will inform what wellness program to create

Reasoning

- **Inductive reasoning** is common in science,
    - where data are collected and tentative models are developed to describe and predict future behavior—until the appearance of anomalous data forces the model to be revised.
- **Deductive reasoning** is common in mathematics and logic,
    - where elaborate structures of irrefutable theorems are built up from a small set of basic axioms and rules.

Reasoning

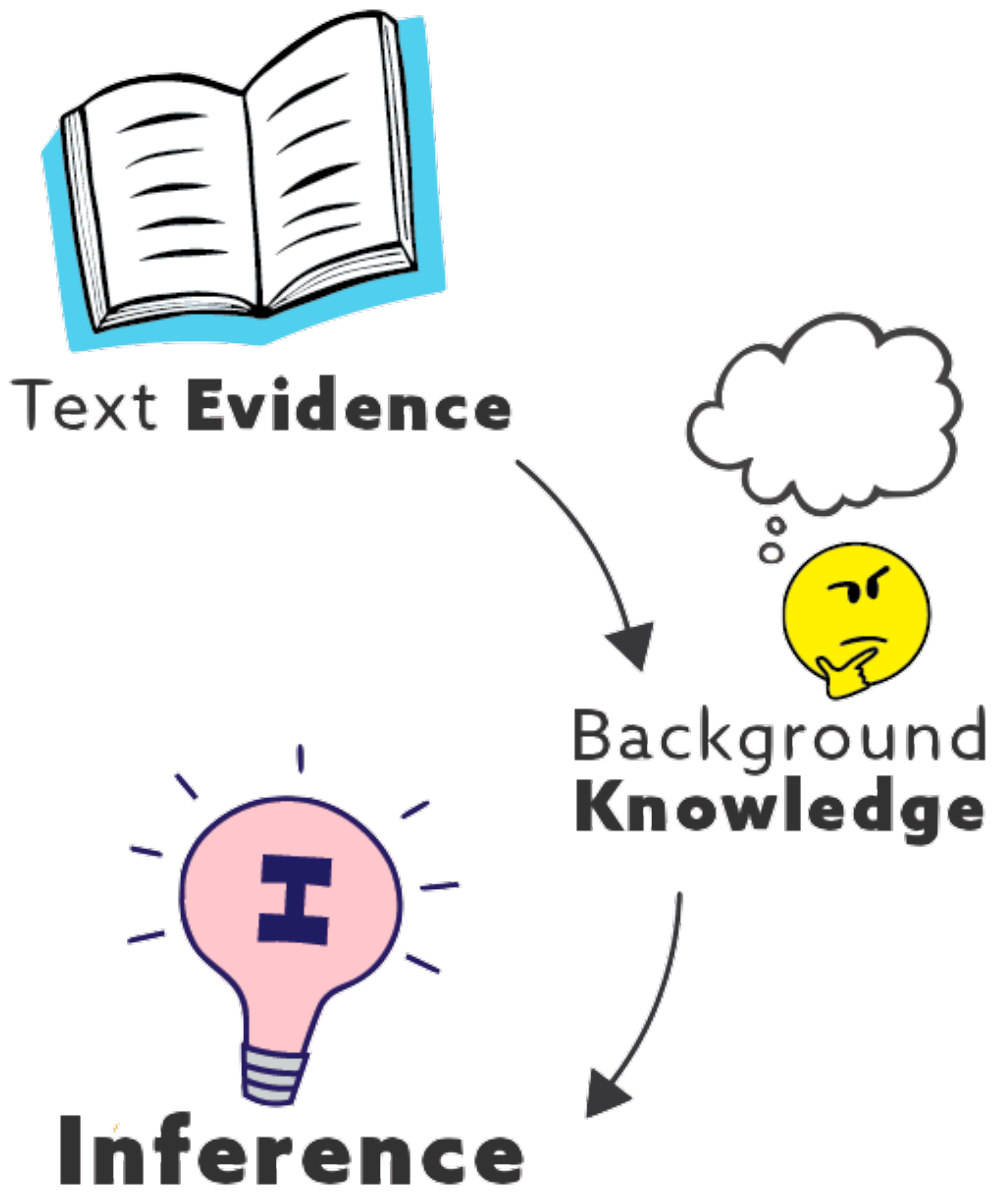
- Other Reasoning methods:
  - Analogical reasoning
  - Abductive reasoning
  - Fallacious reasoning
- Terminologies:
  - **logic**: Step-by-step thinking about how a problem can be solved or a conclusion can be reached.
  - **inference**: A conclusion drawn from true or assumed-true facts.
  - **syllogism**: A type of deductive reasoning, often in the form “All A are B; C is A; therefore, C is B.”
  - **reason**: The capacity for consciously making sense of the world based on logic and evidence.

Reasoning

- To reason is to draw inferences.
- An inference is the process of reasoning from what we think is true to what else is true.
- Important is that an inference is synonymous with the reasoning of an ***argument*** or what we call metaphorically a trail of reasoning.
- Reasoning is the ability to reason (inductive or deductive) and to draw inferences based on situations
- **Reason** is how we form **inferences** about the world

Inference

- Inference is a conclusion reached on the basis of the facts and evidence.
- To reach to a conclusion is the inference process.
- So generating the conclusions from evidence and facts is termed as Inference.
- Application: **Inference Engine**



Inference Engine

- An inference engine is a component of the system that applies logical rules to the knowledge base to deduce new information.
- The first inference engines were components of expert systems.
- The typical expert system consisted of a knowledge base and an inference engine.
- The knowledge base stored facts about the world.
- The inference engine applies logical rules to the knowledge base and deduced new knowledge.
- This process would iterate as each new fact in the knowledge base could trigger additional rules in the inference engine.

Inference Engine

- The logic that an inference engine uses is typically represented as **IF-THEN** rules.
- The general format of such rules is:

  IF <logical expression> THEN <logical expression>

  (what logicians call *modus ponens*)

- A simple example of ***modus ponens*** "If you are human then you are mortal":

  $\text{Human}(x) \Longrightarrow \text{Mortal}\ (x)$

Inference Engine

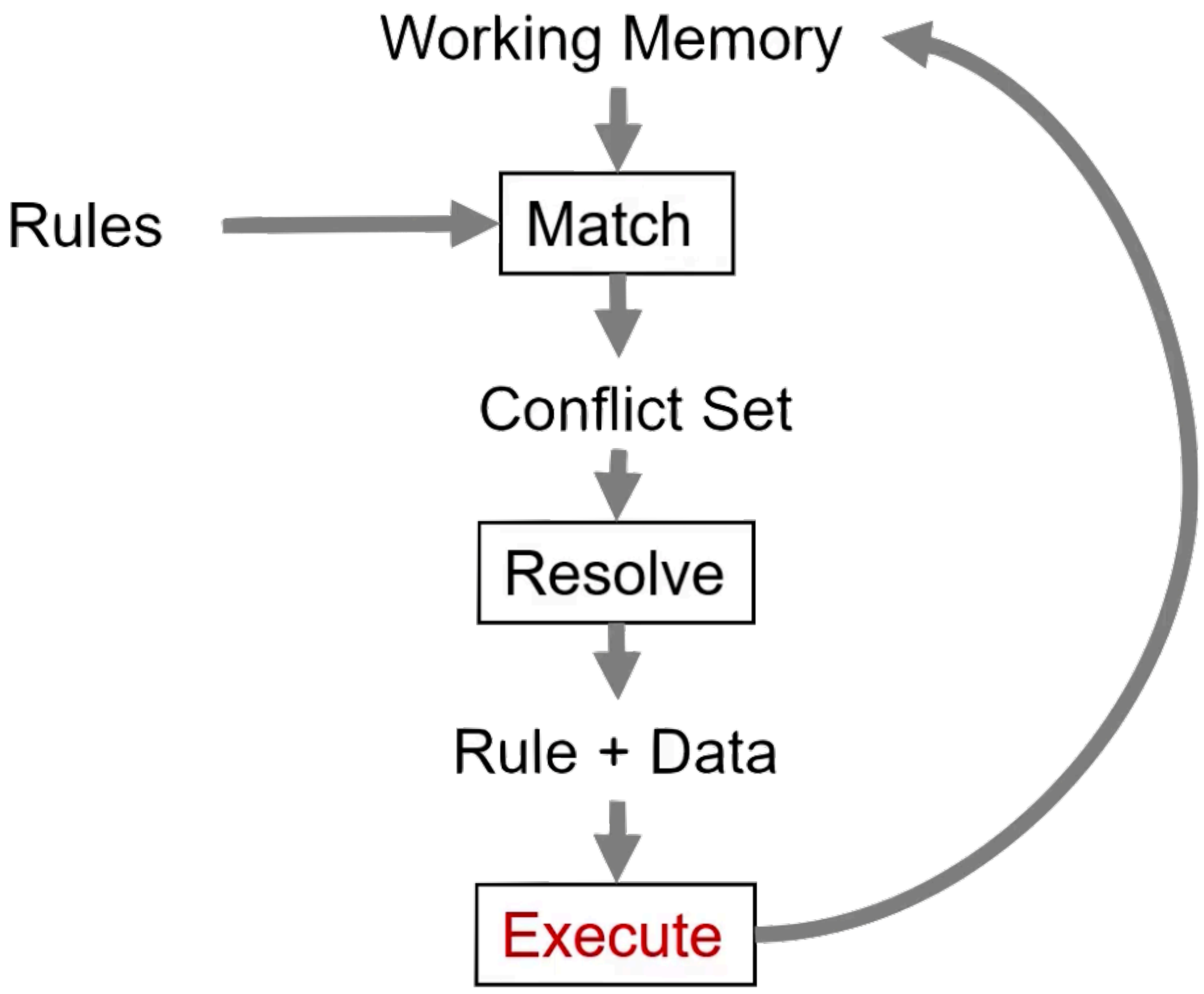
- A knowledge base is an organized collection of facts about the system's domain.
- An inference engine interprets and evaluates the facts in the knowledge base in order to provide an answer.
- The inference engine enables the expert system to draw deductions from the rules in the KB.
- For example,
  - if the KB contains the production rules:
    - "if *x*, then *y*" and "if *y*, then *z*,"
  - the inference engine is able to deduce:
    - "if *x*, then *z*."

Inference Engine

- To recommend a solution, Inference engines work primarily in one of two modes:
  - Forward chaining
    - It starts with the known facts and asserts new facts.
  - Backward chaining
    - It starts with goals, and works backward to determine what facts must be asserted so that the goals can be achieved
- Both backward and forward chaining reasoning progress according to the *modus ponens* form of deductive reasoning.
  - In other words, X implies Y is true. X is true, and therefore Y must be true.

Inference Engine

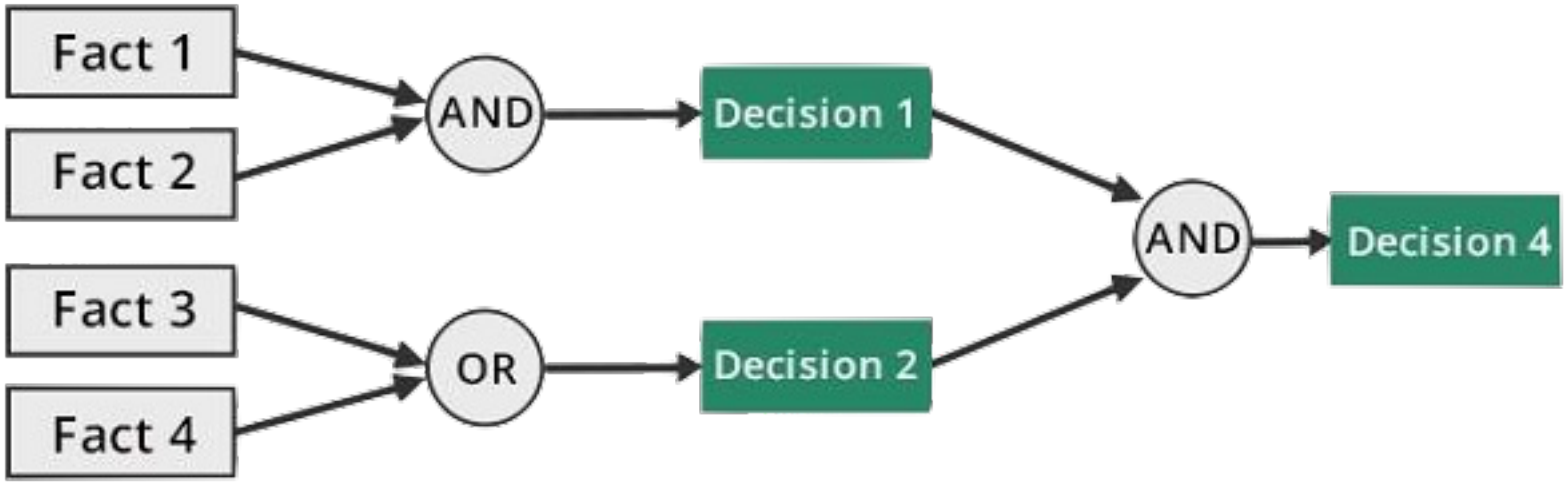
- The actions performed by inference engines tend to progress through three stages:
  - match rules
  - select rules (Resolve)
  - execute rules.
- Match rules is an action in which an inference engine finds all rules triggered by the contents of a knowledge base.
- Select rules is an action which discerns which order rules should be applied in (this will differ for forward or backward chaining, or by other machine learning inputs)
- Execute rules applies rules to existing knowledge through forward or backward chaining
- Once execute rules is completed, match rules is re-started until there are no more opportunities for either forward or backward chaining deductions



Forward Chaining

- It is a strategy of an expert system to answer the question, “What can happen next?”
- Here, the Inference Engine follows the chain of conditions and derivations and finally deduces the outcome.
- It considers all the facts and rules, and sorts them before concluding to a solution.
- This strategy is followed for working on conclusion, result, or effect.
- For example, prediction of share market status as an effect of changes in interest rates.

Forward Chaining

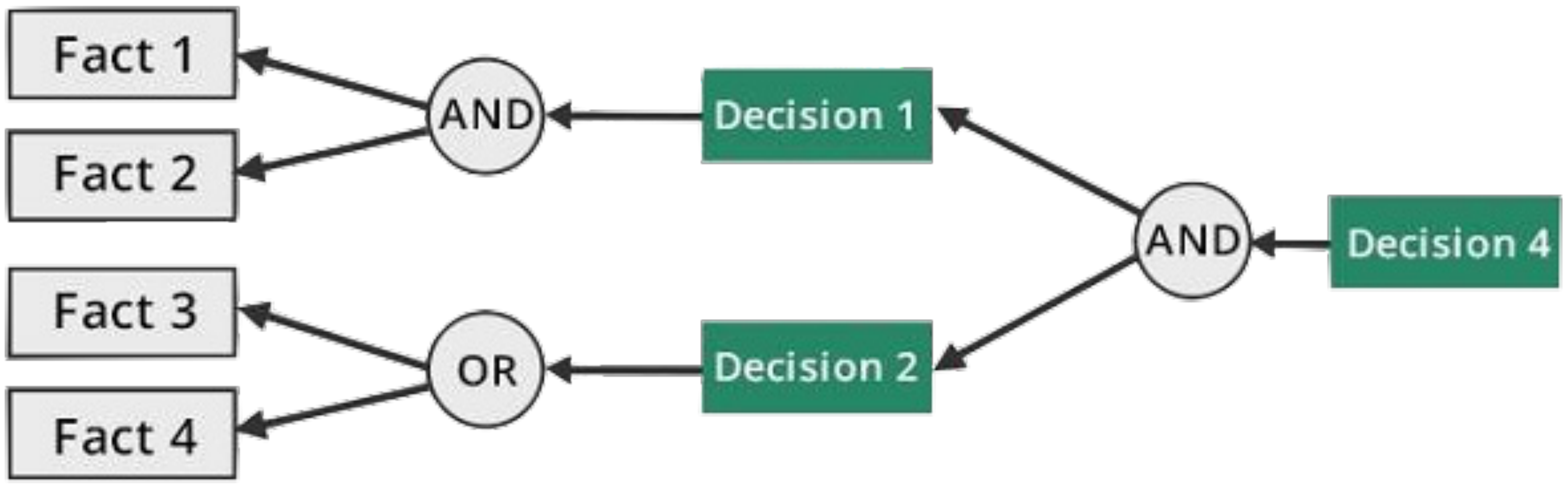


Forward chaining reasoning methods start with available data and utilize rules to infer new data. In short, forward chaining starts with known facts and uses them to create new facts.

Backward Chaining

- With this strategy, an expert system finds out the answer to the question, “Why this happened?”
- On the basis of what has already happened, the Inference Engine tries to find out which conditions could have happened in the past for this result.
- This strategy is followed for finding out cause or reason.
- For example, diagnosis of blood cancer in humans.

Backward Chaining



Backward chaining reasoning methods begin with a list of hypotheses and work backwards to see if data, once plugged into rules support these hypotheses. In short, backward chaining highlights what facts must be true to support a hypothesis.

Backward vs Backward

Forward Chaining	Backward Chaining
Planning monitoring and control	diagnosis
data-driven	goal-driven (hypothesis)
bottom-up reasoning	top-down reasoning
find possible conclusions supp orted by given facts	find facts that support a given hypothesis
Follows: breadth-first search	Follows: depth-first search
antecedents (LHS) control evaluation	consequents (RHS) control evaluation

Backward vs Backward

- **If there is clear hypotheses, then backward chaining is likely to be** better; e.g., Diagnostic problems or classification problems, Medical expert systems

- **Forward chaining may be better if there is less clear hypothesis and** want to see what can be concluded from current situation; e.g., Synthesis systems - design / configuration.

Inference Rules

- Proofs in mathematics are valid arguments that establish the truth of mathematical statements.
  - An **argument** is a sequence of statements that end with a conclusion.
  - The argument is **valid** if the **conclusion** (final statement) follows from the truth of the preceding statements (**premises**).
- **Rules of inference** are templates for building valid arguments.

Inference Rules

- Proof Methods:
  - **Formal Proof:** Proving useful theorems using formal proofs would result in long and tedious proofs, where every single logical step must be provided.
  - **Informal Proof:** Proofs used for human consumption (rather than for automated derivations by the computer) are usually informal proofs, where steps are combined or skipped, axioms or rules of inference are not explicitly provided.

Inference Rules

- A rule of inference (inference rule or transformation rule) is a logical form consisting of a function which
  - takes premises,
  - analyzes their syntax, and
  - returns a conclusion (or conclusions).
- **For example:** modus ponens (a rule of inference)
  - takes two premises, 1) ”If p then q" and 2) ”p", and
  - returns the conclusion "q".

Inference Rules

- In formal logic, rules of inference are usually given in the following standard form:

$$\frac{\begin{array}{l}\text{Premise\#1}\\ \text{Premise\#2}\\ \dots\\ \text{Premise\#n}\end{array}}{\text{Conclusion}}$$

- This expression states that whenever in the course of some logical derivation the given premises have been obtained, the specified conclusion can be taken for granted as well.

Inference Rules

- **For Example:** the ***modus ponens*** rule of propositional logic.
- The Modus Ponens rule is one of the most important rules of inference, and it states that if p and p → q is true, then we can infer that q will be true. It can be represented as:

$$\begin{array}{rl} & p \\ & p \rightarrow q \\ \hline \therefore & q \end{array}$$

- Statement-1: "I am sleepy" (p)
- Statement-2: "If I am sleepy then I go to bed" (P→ q)
- Conclusion: "I go to bed." (q).
- Hence, we can say that, if p→ q is true and p is true then q will be true.

Inference Rules

inference rule	tautology	name
p $p \rightarrow q$ $\therefore\ q$	$(p \land (p \rightarrow q)) \rightarrow q$	Modus ponens (mode that affirms)
$\neg q$ $p \rightarrow q$ $\therefore\ \neg p$	$(\neg q \land (p \rightarrow q)) \rightarrow \neg p$	Modus tollens (mode that denies)
$p \rightarrow q$ $q \rightarrow r$ $\therefore\ p \rightarrow r$	$((p \rightarrow q) \land (q \rightarrow r)) \rightarrow (p \rightarrow r)$	hypothetical syllogism
$p \lor q$ $\neg p$ $\therefore\ q$	$((p \lor q) \land (\neg p)) \rightarrow q$	disjunctive syllogism

Inference Rules

$\begin{array}{rl} & p \\ \hline \therefore & p \lor q \end{array}$	$p \rightarrow (p \lor q)$	addition
$\begin{array}{rl} & p \land q \\ \hline \therefore & p \end{array}$	$(p \land q) \rightarrow p$	simplification
$\begin{array}{rl} & p \\ & q \\ \hline \therefore & p \land q \end{array}$	$((p) \land (q)) \rightarrow (p \land q)$	conjunction
$\begin{array}{rl} & p \lor q \\ & \neg p \lor r \\ \hline \therefore & q \lor r \end{array}$	$((p \lor q) \land (\neg p \lor r)) \rightarrow (q \lor r)$	resolution

Inference Rules

- Rules of Inference for Quantified Statements

Rule of Inference	Name
$\dfrac{\forall x P(x)}{\therefore\ P(c)}$	Universal instantiation
$\dfrac{P(c) \text{ for an arbitrary } c}{\therefore\ \forall x P(x)}$	Universal generalization
$\dfrac{\exists x P(x)}{\therefore\ P(c) \text{ for some element } c}$	Existential instantiation
$\dfrac{P(c) \text{ for some element } c}{\therefore\ \exists x P(x)}$	Existencial generalization

Inference Rules

- Combining Rules of Inference for Propositions and Quantified Statements

Universal Modus Ponens

$$\begin{array}{rl} & \forall x(P(x) \rightarrow Q(x)) \\ & P(a), \text{ where } a \text{ is a particular element in the domain} \\ \hline \therefore & Q(a) \end{array}$$

Universal Modus Tollens

$$\begin{array}{rl} & \forall x(P(x) \rightarrow Q(x)) \\ & \neg Q(a), \text{ where } a \text{ is a particular element in the domain} \\ \hline \therefore & \neg P(a) \end{array}$$

Knowledge-based Agent

- The **problem-solving agents** know things, but only in a very limited, inflexible sense.
- Intelligent agents need knowledge about the world in order to reach good decisions.
- A **knowledge-based agent** is composed of
  - a knowledge base and
  - an inference mechanism.
- The **knowledge-based agent** operates
  - by storing sentences about the world in its knowledge base,
  - using the inference mechanism to infer new sentences, and using these sentences to decide what action to take.

Uncertainty

- You have learned knowledge representation using first-order logic and propositional logic with certainty, which means we were sure about the predicates.
- With this knowledge representation, we might write $A \rightarrow B$, which means if A is true then B is true
- Consider a situation where we are not sure about whether A is true or not then we cannot express this statement, this situation is called **uncertainty**.

Uncertainty

- In real life, it is not always possible to determine the state of the environment as it might not be clear.
- Due to partially observable or non-deterministic environments, **agents** may need to handle **uncertainty** and deal with.
- For example: Diagnosis—whether for medicine, automobile repair, or whatever—almost always involves **uncertainty**.

Uncertain Reasoning: Example

- **Diagnosing a dental patient's toothache:**
  - Consider a simple rule (using propositional logic):

    Toothache → Cavity.

  - This rule is not complete because not all patients with toothaches have cavities; some of them have gum disease, an abscess, or one of several other problems:

    Toothache → Cavity ∨ GumProblem ∨ Abscess . . .

  - Unfortunately, in order to make the rule true, we have to add an almost unlimited list of possible problems. We could try turning the rule into a causal rule:

    Cavity → Toothache.

    But this rule is not right either; not all cavities cause pain. make it logically exhaustive: to augment the left-hand side with all the qualifications required for a cavity to cause a toothache.

  - The only way to fix the rule is to make this rule complete by list all the possible causes of toothache. But this is not feasible (due to Laziness, theoretical ignorance and practical ignorance).

Logic fails due to uncertainty

- Trying to use logic to cope with a domain like medical diagnosis thus fails for three main reasons:
  - **Laziness**: It is too much work to list the complete set of antecedents or consequents needed to ensure an exceptionless rule and too hard to use such rules.
  - **Theoretical ignorance:** Medical science has no complete theory for the domain.
  - **Practical ignorance:** Even if we know all the rules, we might be uncertain about a particular patient because not all the necessary tests have been or can be run.

Probability Theory

- The connection between ***toothaches*** and ***cavities*** is just not a logical consequence in either direction.
- This is typical of the medical domain, as well as most other judgmental domains: law, business, design, automobile repair, gardening, dating, and so on.
- The agent's knowledge can provide only a **degree of belief** in the relevant sentences.
- A tool for dealing with ***degrees of belief*** is ***probability theory***.
- We use probability theory for reasoning on uncertain knowledge

Probability Theory

- Variables in probability theory are called **random variables. (**begin with an uppercase letter **)**
- Every random variable has a domain—the set of possible values it can take on. (written in lower case)
  - The domain for two dice is the set {2,...,12} and
  - The domain of Die 1 is {1,...,6}.
  - A Boolean random variable has the domain {true,false}
- 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:
  - $P\ (cavity \mid \neg toothache \wedge teen\ ) = 0.1$.

Probability Theory

- **Probability** can be defined as a chance that an uncertain event will occur.
- **Probability** 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.
- $0 \leq P(A) \leq 1$, where P(A) is the probability of an event A. (P(A) = Degree of belief)
- $P(A) = 0$, indicates **total uncertainty** in an event A.
- $P(A) = 1$, indicates **total certainty** in an event A.

Probability Theory

- The **probability** of an uncertain event A, denoted as P(A) is given by

$$P(A) = \frac{\text{No. of event A occurs}}{\text{Total Number of trials/outcomes}}$$

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

Probability Theory

- **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.

- **Marginal probability** of a subset of a collection of random variables is the probability distribution of the variables contained in the subset. It gives the probabilities of various values of the variables in the subset without reference to the values of the other variables. It is the probability of an event irrespective of the outcomes of other random variables, e.g. P(A).

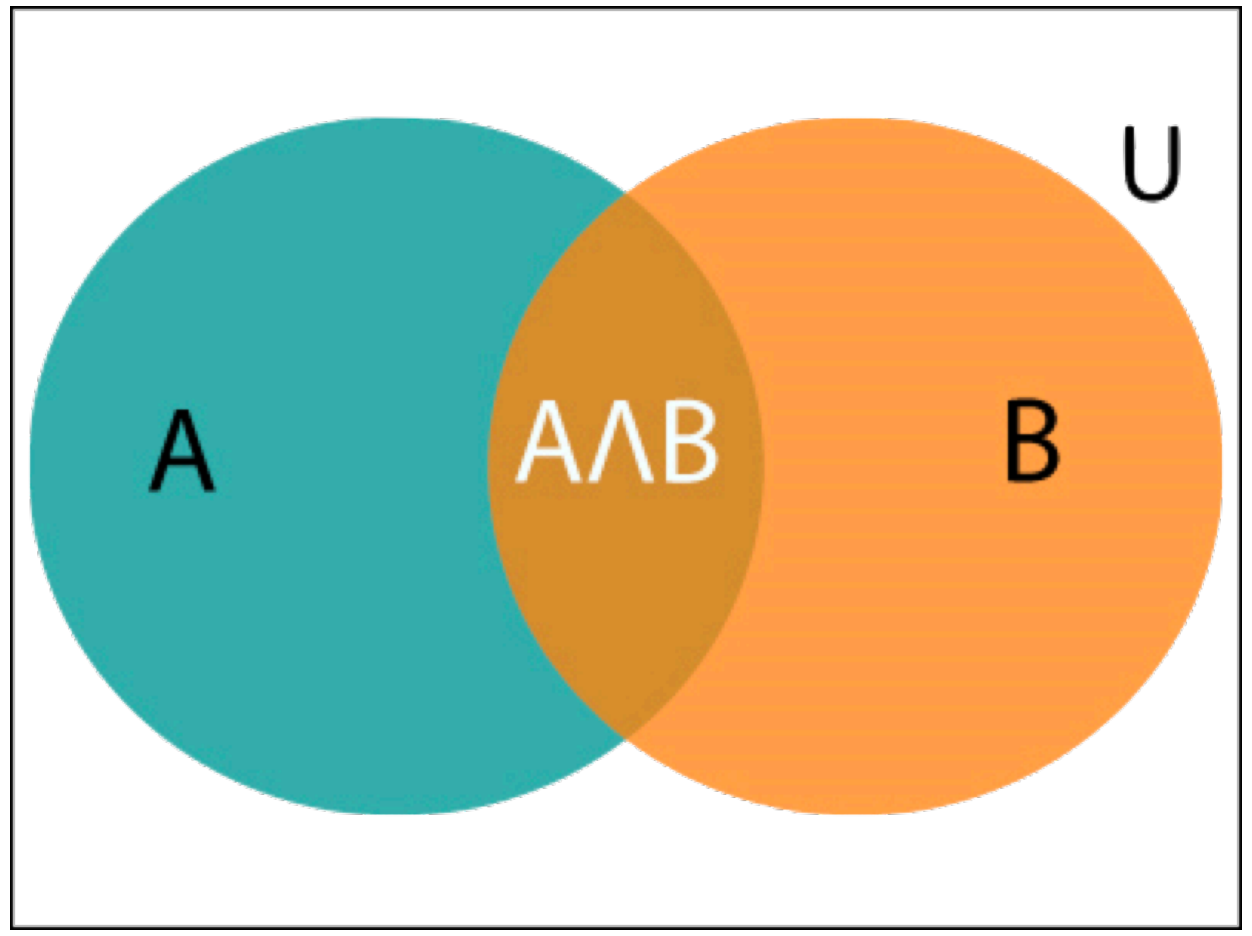
- **Joint Probability**: Probability of two (or more) simultaneous events, e.g. P(A and B) or P(A, B). The joint probability is the probability of two (or more) simultaneous events, often described in terms of events A and B from two dependent random variables, e.g. X and Y. The joint probability is often summarized as just the outcomes, e.g. A and B.

Probability Theory

- **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, "the probability of A under the conditions of B", it can be written as:

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



- Where $P(A \wedge B)$ = Joint probability of A and B

- $P(B)$ = Marginal probability of B.

Probability Theory

- **Conditional probability: 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?

- **Solution**: Let, A is an event that a student likes Mathematics

  B is an event that a student likes English.

  $P(A \wedge B)$ = 40% = 0.4

  $P(B)$ = 70% = 0.7

$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)} = \frac{0.4}{0.7} = 57\,\%$$

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

Bayes' Theorem

- also known as Bayes' rule, Bayes' law, or Bayesian reasoning, which determines the probability of an event with uncertain knowledge.

- Bayes' theorem was named after the British mathematician Thomas Bayes. The Bayesian inference is an application of Bayes' theorem, which is fundamental to Bayesian statistics.

- Bayes' theorem describes the probability of an event, based on prior knowledge of conditions that might be related to the event

- It relates the conditional probability and marginal probability and prior probability of two random events.

Bayes' Theorem

- Bayes' theorem is stated mathematically as:

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

- Where A and B are events and $P(B) \neq 0$.
  - $P(A \mid B)$ is a conditional probability: the probability of event A occurring given that B is true. It is also called the posterior probability of A given B.
  - $P(B \mid A)$ is also a conditional probability: the probability of event B occurring given that A is true.
  - $P(A)$ is called the **prior probability**, probability of hypothesis before considering the evidence and $P(B)$ is marginal probability. These are the probabilities without any given conditions.
  - A and B must be different events.
  - This simple equation underlies most modern AI systems for probabilistic inference.

Bayes' Theorem

- Suppose we want to perceive the effect of some unknown cause, and want to compute that cause, then the Bayes' rule becomes:

$$P(cause \mid effect) = \frac{P(effect \mid cause)P(cause)}{P(effect)}$$

Bayes' Theorem

- **Example 1:** A doctor is aware that disease meningitis causes a patient to have a stiff neck, and it occurs 80% of the time. He is also aware of some more facts, which are given as follows:

  - The Known probability that a patient has meningitis disease is 1/30,000.

  - The Known probability that a patient has a stiff neck is 2%

- what is the probability that a patient has diseases meningitis with a stiff neck?

Bayes' Theorem

- **Example 1: Solution:**

Let A be the proposition that patient has stiff neck (effect) and B be the proposition that patient has meningitis (cause). So we can calculate the following as:

$P(A \mid B) = 0.8$

$P(B) = 1/30000$

$P(A) = 0.02$

The probability that a patient has diseases meningitis with a stiff neck is given by:

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)} = \frac{0.8\frac{1}{30000}}{0.02} = \frac{1}{750} = 0.00133333$$

Hence, we can assume that 1 patient out of 750 patients has meningitis disease with a stiff neck.

Applications of Bayes' Theorem

- Naive Bayes' Classifiers
- Discriminant Functions and Decision Surfaces
- Bayesian Parameter Estimation

Bayesian Network

- A Bayesian network (BN) is a probabilistic graphical model for representing knowledge about an uncertain domain
- It represents a set of random variables and their conditional dependencies via a directed acyclic graph (DAG)
- In BN, each node corresponds to a random variable and each edge represents the conditional probability for the corresponding random variables.
- BNs are also called belief networks or Bayes nets.
- Due to dependencies and conditional probabilities, a BN corresponds to a directed acyclic graph (DAG) where no loop or self connection is allowed.
- It measures the conditional dependence structure of a set of random variables based on the Bayes theorem:

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

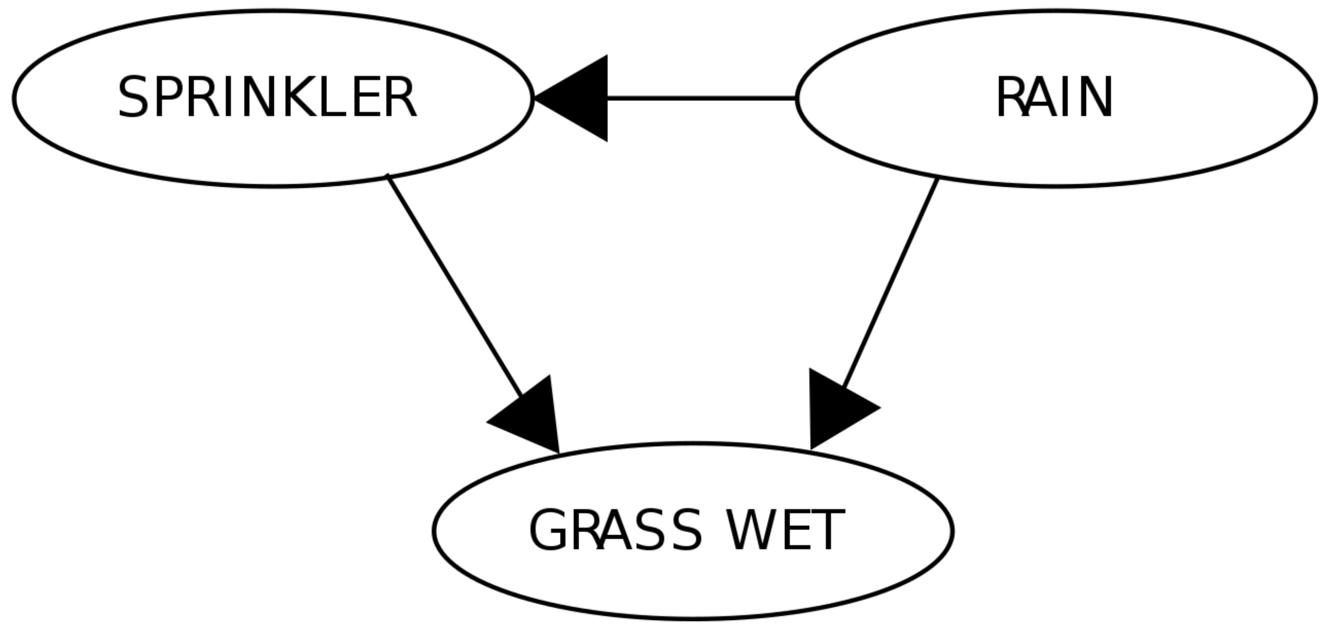
Bayesian Network

- A Bayesian network could represent the probabilistic relationships between diseases and symptoms.
- Given symptoms, the network can be used to compute the probabilities of the presence of various diseases.

Bayesian Network

- Example 1:
- Two events can cause grass to be wet: an active sprinkler or rain.
- Rain has a direct effect on the use of the sprinkler (namely that when it rains, the sprinkler usually is not active).
- This situation can be modeled with a Bayesian network (shown to the right).
- Each variable has two possible values, T (for true) and F (for false).

	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



RAIN	
T	F
0.2	0.8

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

Bayesian Network

- Example 1:

- The joint probability function is, by the chain rule of probability,

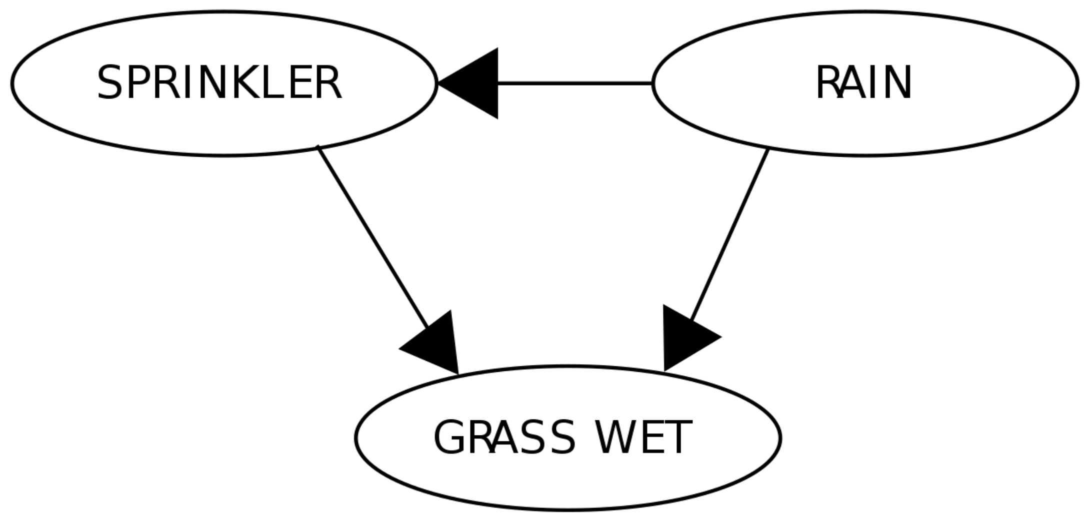
$$\Pr(G,S,R) = \Pr(G \mid S,R)\Pr(S \mid R)\Pr(R)$$

- where G = "Grass wet (true/false)", S = "Sprinkler turned on (true/false)", and R = "Raining (true/false)".

- The model can answer questions: "What is the probability that it is raining, given the grass is wet?" by using the conditional probability formula and summing over all nuisance variables:

$$\Pr(R=T \mid G=T) = \frac{\Pr(G=T,R=T)}{\Pr(G=T)} = \frac{\sum_{x\in\{T,F\}}\Pr(G=T,S=x,R=T)}{\sum_{x,y\in\{T,F\}}\Pr(G=T,S=x,R=y)}$$

	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



RAIN	
T	F
0.2	0.8

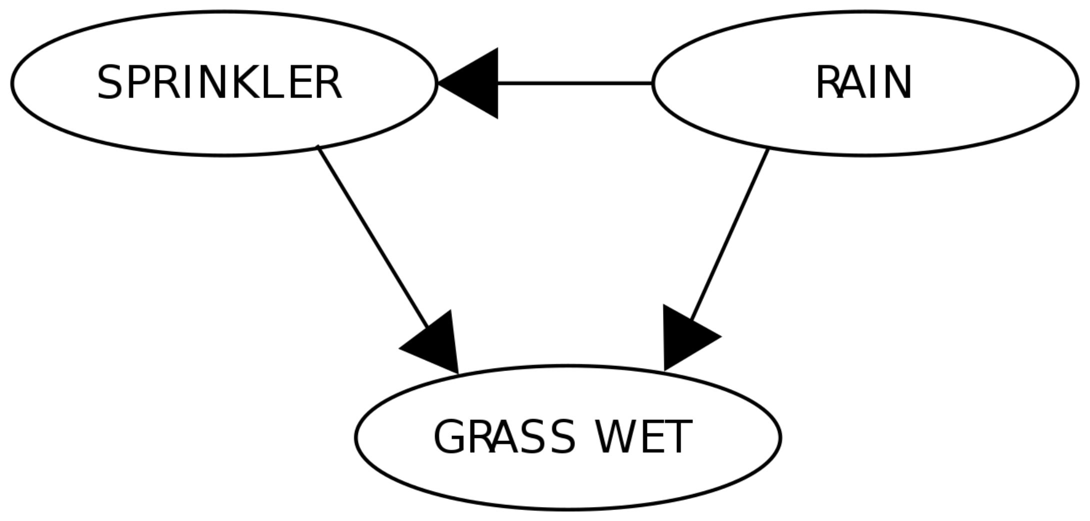
		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

Bayesian Network

- Example 1:

- Using the expansion for the joint probability function $\Pr(G, S, R)$ and the conditional probabilities from the conditional probability tables (CPTs) stated in the diagram, one can evaluate each term in the sums in the numerator and denominator. For example,

	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



RAIN	
T	F
0.2	0.8

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

$$\begin{aligned}
\Pr(G=T, S=T, R=T) &= \Pr(G=T \mid S=T, R=T)\Pr(S=T \mid R=T)\Pr(R=T) \\
&= 0.99 \times 0.01 \times 0.2 \\
&= 0.00198.
\end{aligned}$$

Bayesian Network

- In practice, it is usually possible to obtain only the reversed conditional probability, i.e. probability of the evidence given the cause, the probability of observing symptoms if the patient has the disease.
- A Bayesian approach is appropriate in these cases
- Bayesian networks are very useful tools for dealing with
    - uncertainty
    - complexity and
    - causality (relationship between cause and effect)

Bayesian Network

- A Bayesian network represents the causal probabilistic relationship among a set of random variables, their conditional dependences, and it provides a compact representation of a joint probability distribution

- It consists of two major parts:

1. **a directed acyclic graph:** is a set of random variables represented by nodes. If there exists a causal probabilistic dependence between two random variables in the graph, the corresponding two nodes are connected by a directed edge. the directed edge from a node A to a node B indicates that the random variable A causes the random variable B. Since the directed edges represent a static causal probabilistic dependence, cycles are not allowed in the graph

2. **a set of conditional probability distributions:** A conditional probability distribution is defined for each node in the graph. In other words, the conditional probability distribution of a node (random variable) is defined for every possible outcome of the preceding causal node(s).

Case-based Reasoning

- Case-based reasoning means
  - using old experiences to understand and solve new problems.
- In case-based reasoning, a reasoner remembers a previous situation similar to the current one and uses that to solve the new problem.
- Case- based reasoning can mean adapting old solutions to meet new demands;
  - using old cases to explain new situations;
  - using old cases to critique new solutions; or
  - reasoning from precedents to interpret a new situation (much like lawyers do) or create an equitable solution to a new problem (much like labor mediators do).

Case-based Reasoning

- If we watch the way people around us solve problems, we are likely to observe case-based reasoning in use all around us.
    - Attorneys are taught to use cases as precedents for constructing and justifying arguments in new cases.
    - Mediators and arbitrators are taught to do the same.
- In general, the second time solving some problem or doing some task is easier than the first because we remember and repeat the previous solution.
    - We are more competent the second time because we remember our mistakes and go out of our way to avoid them.

Case-based Reasoning

- Consider, for example, a doctor faced with a patient who has an unusual combination of symptoms.
    - If he's seen a patient with similar symptoms previously, he is likely to remember the old case and propose the old diagnosis as a solution to his new problem.
    - If proposing those disorders was time-consuming previously, this is a big savings of time.
    - Of course, the doctor can't assume the old answer is correct. He/she must still validate it for the new case in a way that doesn't prohibit considering other likely diagnoses.
    - Nevertheless, remembering the old case allows him to generate a probable answer easily.

Case-based Reasoning

- The quality of a case-based reasoner's solutions depends on four things:
  - the experiences it's had:
    - The less experienced reasoner will always have fewer experiences to work with than the more experienced one.
    - less experienced reasoner won't necessarily be worse
    - Those experiences (cases) should cover the goals and subgoals that arise in reasoning and should include both successful and failed attempts at achieving those goals. Successful attempts will be used to propose solutions to new problems. Failed attempts will be used to warn of the potential for failure.
  - its ability to understand new situations in terms of those old experiences,
  - its skill at adaptation:
    - adaptation, is the process of fixing up an old solution to meet the demands of the new situation.
  - its skill at evaluation.

Case-based Reasoning

- it is ability to understand new situations in terms of those old experiences:
- It has two parts:
  - ***recalling*** (called *indexing problem*) old experiences and
  - ***interpreting*** the new situation in terms of the recalled experiences.

Case-based Reasoning

- ***Recalling*** *(indexing problem):*
  - finding in memory the experience closest to a new situation. That is it is a task of assigning indexes to experiences stored in memory so that they can be recalled under appropriate circumstances.
  - Recalling cases appropriately is at the core of case-based reasoning.

Case-based Reasoning

- ***Interpretation***:
    - Interpretation is the process of comparing the new situation to recalled experiences.
    - When problem situations are interpreted, they are compared and contrasted to old problem situations.
    - The result is an interpretation of the new situation, the addition of inferred knowledge about the new situation, or a classification of the situation.

Case-based Reasoning

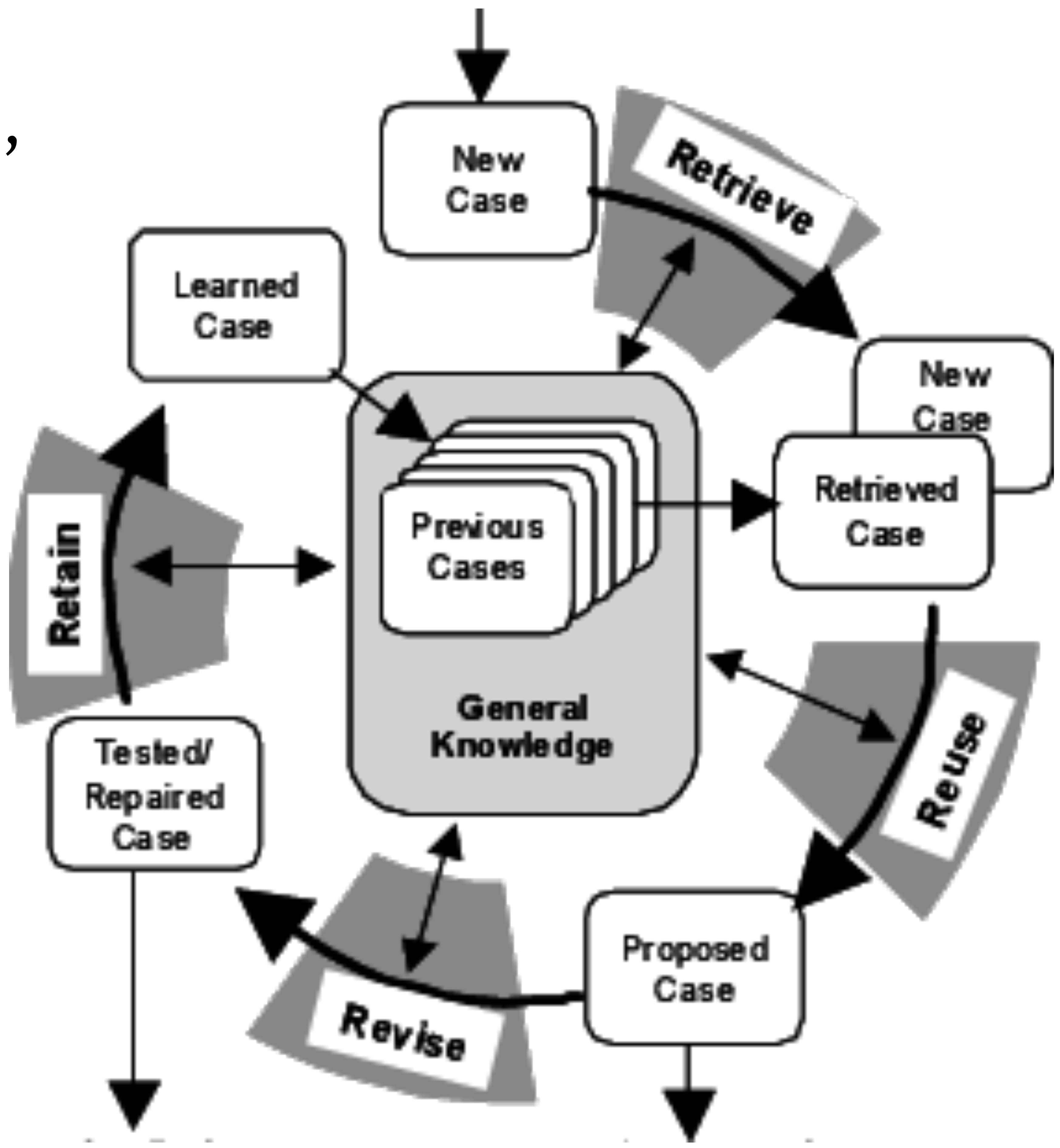
- its skill at adaptation:
    - adaptation, is the process of fixing up an old solution to meet the demands of the new situation.
    - Adaptation inserts something new into an old solution, to delete something, 6r to make a substitution.

Case-based Reasoning

- its skill at evaluation:
  - One of the hallmarks of a case-based reasoner is its ability to learn from its experiences, as a doctor might do when he caches a hard-to-solve problem so that he can solve it easily another time.
  - In order to learn from experience, a reasoner requires feedback so that it can interpret what was right and wrong with its solutions.
  - Without feedback, the reasoner might get faster at solving problems but would repeat its mistakes and never increase its capabilities.
  - Thus, ***evaluation*** and consequent ***repair*** are important contributors to the expertise of a case-based reasoner.
  - Evaluation can be done in the context of the outcomes of other similar cases, can be based on ***feedback*** or can be based on simulation.

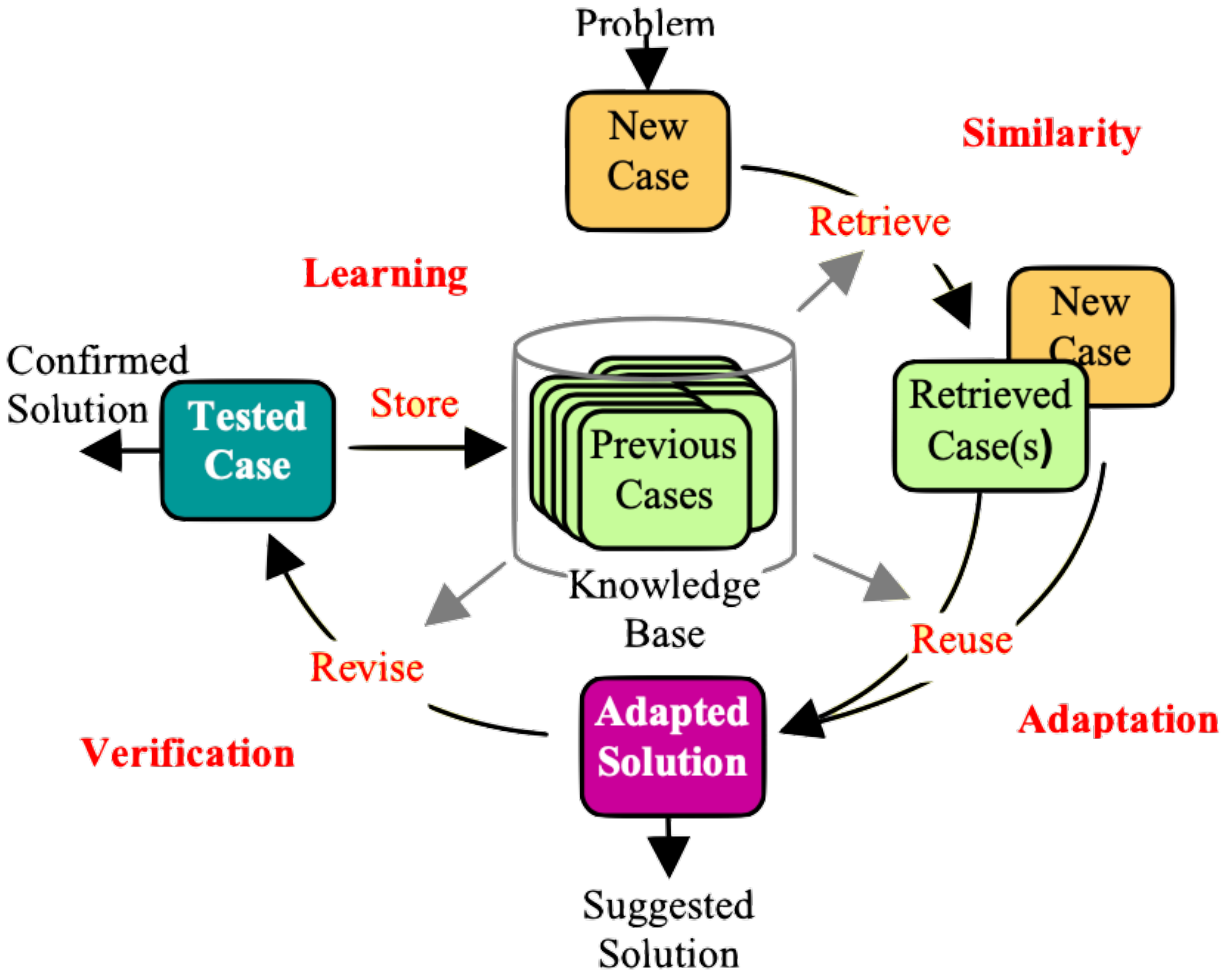
Case-based Reasoning

- Case-Based Reasoning (CBR) is used to model, store, and re-use historical data, and capture knowledge for problem-solving tasks.
- An important feature of CBR is the ability to learn from past cases/ situations.
- A CBR system stores and organises past situations, then chooses situations similar to the problem at hand and adapts a solution based on the previous cases.
- An overview of the CBR process is illustrated in Figure



Case-based Reasoning

- Case-based reasoning has been formalized for purposes of computer reasoning as a four-step process:



Case-based Reasoning

- **Retrieve:** Given a target problem, retrieve cases relevant to solving it from memory. A case consists of a problem, its solution, and, typically, annotations about how the solution was derived.
- **Reuse:** Map the solution from the previous case to the target problem. This may involve adapting the solution as needed to fit the new situation.
- **Revise:** Having mapped the previous solution to the target situation, test the new solution in the real world (or a simulation) and, if necessary, revise. (Suggesting a solution based on the experience and adapting it to meet the demands of the new situation.)
- **Retain:** After the solution has been successfully adapted to the target problem, store the resulting experience as a new case in memory.

Case-based Reasoning

- Case-based reasoning is a prominent type of analogy solution making.
- case-based reasoning is not only a powerful method for computer reasoning, but also a pervasive behavior in everyday human problem solving; or, more radically, that all reasoning is based on past cases personally experienced.
- Principles from CBR research serve as a foundation for applied computer systems for tasks such as
  - supporting human decision-making,
  - aiding human learning, and
  - facilitating access to electronic information repositories.

Case-based vs Knowledge-based

- Case-based reasoning has several differences from other AI approaches, such as knowledge-based systems (KBS). Rather than relying completely on general knowledge of a problem domain or making associations along generalized relationships between problem descriptors and conclusions, CBR employs the specific knowledge of previously experienced, concrete problem situations. CBR also offers incremental, sustained learning in that each time a problem is solved a new experience is retained and can be applied for future problems.
- One of the key differences between rule-based and case-based knowledge engineering is that automatic case-indexing techniques drastically reduce the need to extract and structure specific rule-like knowledge from an expert – the most time-consuming part of rule-based knowledge engineering.

Case-based Reasoning: Advantages

- remembering past experiences helps learners avoid repeating previous mistakes, and the reasoner can discern what features of a problem are significant and focus on them.
- it reflects how people work. Because no knowledge must be elicited to create rules or methods, development is easier. Another benefit is that systems learn by acquiring new cases through use, and this makes maintenance easier.
- CBR also enables the reasoner to propose solutions to problems quickly.
- **Disadvantage:** On the negative side, critics claim that the main premise of CBR is based on anecdotal evidence and that adapting the elements of one case to another may be complex and potentially lead to inaccuracies.

Case-based Reasoning

- Case-based reasoning is used for classification and for regression.
- If the cases are simple, one algorithm that works well is to use the ***k-nearest neighbors*** for some given number ***k*** .
- Given a new example, the k training examples that have the input features closest to that example are used to predict the target value for the new example. The prediction could be the mode, average, or some interpolation between the prediction of these k training examples, perhaps weighting closer examples more than distant examples.
- For this method to work, a distance metric is required that measures the closeness of two examples. First define a metric for the domain of each feature, in which the values of the features are converted to a numerical scale that is used to compare values.

THANK YOU

End of Chapter

https://www.slideshare.net/DigiGurukulBlog/artificial-intelligence-notes-unit-1