MCA

CHAPTER 5

KNOWLEDGE REPRESENTATION

5.1 INTRODUCTION

One definition of Artificial Intelligence (AI) is that it is the study of intelligent behaviour achieved through computational means. Knowledge representation and reasoning, is that part of AI that is concerned with how an agent uses what it knows in deciding what to do. It is the study of thinking as a computational process. The knowledge can be represented either in the form of facts or in some formalism. Representation of knowledge and the reasoning process are central to the entire field of artificial intelligence. Knowledge must be meaningfully represented so that we know how it relates back to the real world. A knowledge representation scheme provides a mapping from features of the world to a formal language.

If we have knowledge that is sufficient to solve a problem, we have to search our goal in that knowledge. The primary component of a knowledge-based agent is its *knowledge-base*. A knowledge-base is a set of sentences. Each sentence is expressed in a language called the *knowledge representation language*. Sentences represent some assertions about the world. There must mechanisms to derive new sentences from old ones. This process is known as *inferencing* or *reasoning*. Inference must obey the primary requirement that the new sentences should follow logically from the previous ones.

Knowledge is subject of proposition

5.2 WHAT IS KNOWLEDGE ?

This is a question that has been discussed by philosophers since the ancient Greeks, and it is still not totally demystified.

Knowledge is an abstract term that attempts to capture an individual's understanding of a given subject. Knowledge is a subset of *information*. But it is a subset that has been extracted, filtered, or formatted in a very special way. More specifically, the information we call knowledge is information that has been subjected to, and passed tests of validation. Common sense knowledge is information that has been validated by common sense experience. Scientific knowledge is information (hypotheses and theories) validated by the rules and tests applied to it by some scientific community. Organizational knowledge in terms of this

framework is information validated by the rules and tests of the organization seeking knowledge. The quality of its knowledge then will be largely dependent on the tendency of its validation rules and tests to produce knowledge that improves organizational performance (the organization's version of objective knowledge).

Knowledge is a relation between a knower and a *proposition*, that is, the idea expressed by a simple declarative sentence.

5.3 TYPES OF KNOWLEDGE

The types of knowledge include procedural knowledge, declarative knowledge and heuristic knowledge.

- **Procedural Knowledge**

  Procedural knowledge is compiled or processed form of information. Procedural knowledge is related to the performance of some task. For example, sequence of steps to solve a problem is procedural knowledge.

- **Declarative Knowledge**

  Declarative knowledge is passive knowledge in the form of statements of facts about the world. For example, mark statement of a student is declarative knowledge.

- **Heuristic Knowledge**

  Heuristics knowledge are rules of thumb or tricks. Heuristic knowledge is used to make judgments and also to simplify solution of problems. It is acquired through experience. An expert uses his knowledge that he has gathered due to his experience and learning.

5.4 WHAT IS KNOWLEDGE REPRESENTATION ?

The concept of *representation* is as philosophically vexing as that of knowledge. Very roughly speaking, representation is a relationship between two domains, where the first is meant to "stand for" or take the place of the second. Usually, the first domain, the representor, is more concrete, immediate, or accessible in some way than the second. For example, the drawing of a circle with a plus below it might stand for the much more abstract concept of womanhood; an elected legislator might stand for his or her constituency etc. The type of representor that we will be most concerned with here is the formal *symbol*, that is, a character or group of characters taken from some predetermined alphabet.

Knowledge is the information about a domain that can be used to solve problems in that domain. To solve many problems requires much knowledge, and this knowledge must be represented in the computer. As part of designing a program to solve problems, we must define how the knowledge will be represented.

Knowledge representation is the field of study concerned with using formal symbols to represent a collection of propositions believed by some supposed agent. Knowledge

representation is the study of ways of picturization of knowledge and how effectively it resembles the representation of knowledge in human brain. Knowledge representation can be defined as *"Knowledge consists of facts, concepts, rules and so on. It can be represented in different forms as mental images in one's thought, as spoken or written words in some language, as graphical or other pictures and as characters, strings or collection of magnetic spots stored in a computer"*.

We argue that the notion of knowledge representation can best be understood in terms of five distinct roles it plays, each crucial to the task at hand :

- A Knowledge Representation (KR) is most fundamentally a surrogate, a substitute for the thing itself, used to enable an entity to determine consequences by thinking rather than acting, *i.e.*, by reasoning about the world rather than taking action in it.
- It is a set of ontological commitments, *i.e.*, an answer to the question: In what terms should I think about the world ?
- It is a fragmentary theory of intelligent reasoning, expressed in terms of three components : (*i*) the representation's fundamental conception of intelligent reasoning ; (*ii*) the set of inferences the representation sanctions ; and (*iii*) the set of inferences it recommends.
- It is a medium for pragmatically efficient computation, *i.e.*, the computational environment in which thinking is accomplished. One contribution to this pragmatic efficiency is supplied by the guidance a representation provides for organizing information so as to facilitate making the recommended inferences.
- It is a medium of human expression, *i.e.*, a language in which we say things about the world.

Knowledge Acquisition

Knowledge Acquisition is the process of acquiring knowledge from a human expert for an expert system, which must be carefully organized into IF-THEN rules or some other form of knowledge representation. Knowledge acquisition is the process of absorbing and storing new information in memory, the success of which is determined by how well the information depends heavily on the representation and organization of the information.

5.5 NEED OF KNOWLEDGE REPRESENTATION

For solving complex problems one needs large amount of knowledge and some mechanisms for manipulating that knowledge to create solutions to new problems. One way to manipulate that knowledge is *knowledge representation*.

Solving problems in a particular domain generally requires knowledge of the objects in the domain and knowledge of how to reason in that domain – both these types of knowledge must be represented. Knowledge must be represented efficiently, and in a meaningful way. Efficiency is important, as it would be impossible (or at least impractical) to explicitly represent every fact that you might ever need. There are just so many potentially useful facts,

most of which you would never even think of. You have to be able to infer new facts from your existing knowledge, as and when needed, and capture general abstractions, which represent general features of sets of objects in the world.

In solving problems in AI we must represent knowledge. There are *two* entities to deal with :

1. Facts

Truth about the real world. These are the things we want to represent. This can be regarded as *knowledge level*.

2. Representation of the Facts

Representations of facts in some chosen formalism. These are the things with which we can actually manipulate data. This can be regarded as the *symbol level* since we usually define the representation in terms of symbols that can be manipulated by programs.

We can structure these entities at *two* levels :

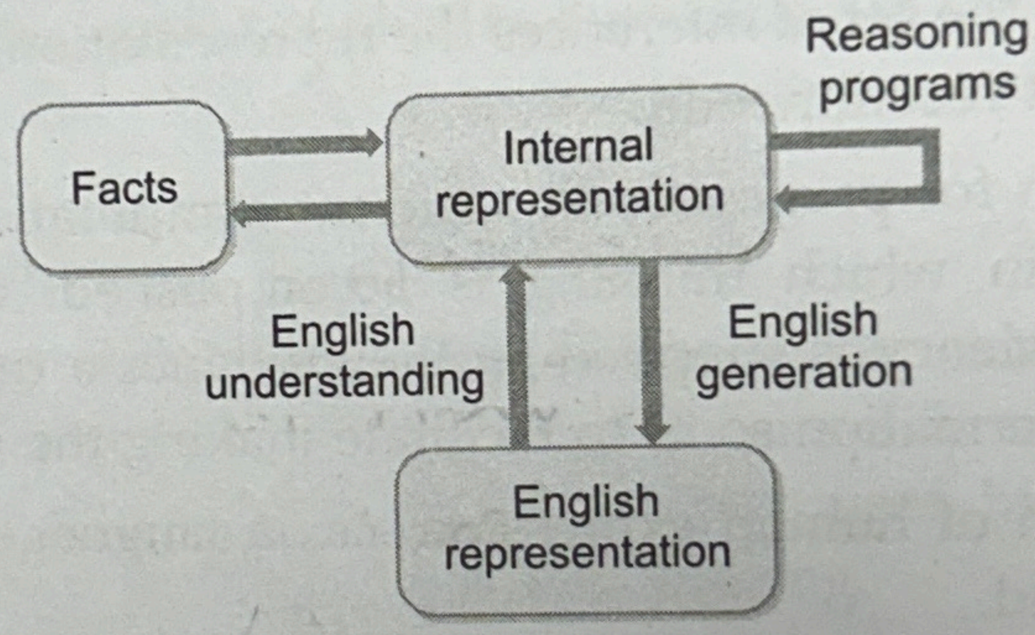

Figure 5.1

(*i*) *Knowledge Level*. At this level the facts are described.

(*ii*) *Symbol Level*. At this level representations of objects are defined in terms of symbols that can be manipulated in programs.

Thus, knowledge is a collection of "facts" from some domain. We need a representation of "facts" that can be manipulated by a program. Normal English is insufficient, too hard currently for a computer program to draw inferences in natural languages. Thus some symbolic representation is necessary. Therefore, we must be able to map "facts to symbols" and "symbols to facts" using forward and backward representation mapping.

5.6 TERMS USED IN KNOWLEDGE REPRESENTATION

Objects	Objects which are entities in the world of agent about which information needs to store.
Events	Events are the actions that occur in agent world.

Performance Performance is the behaviour of agent.

Meta-language Knowledge about what we know. It is a catalogue of items about which agent has knowledge.

Facts Facts are the truth about real world and what we represent. This can be regarded as knowledge level.

The symbol level Symbol level is the limit at which representations of objects are defined in terms of symbol that can be manipulated in programs.

Entailment It is logical association between two sentences. It will be useful for reasoning.

Inference algorithm It is also called as *model checking*. Using this algorithm inference can be drawn and these are stored in knowledge base. An inference algorithm is complete if it can derive any sentence that is entailed. An inference algorithm that derives only entailed sentences is called *sound* or *truth preserving*.

5.7 KNOWLEDGE REPRESENTATION SCHEMES

There are *four* schemes of knowledge representation :

Logical schemes • Predicate calculus • Propositional calculus	**Procedural schemes** • IF..THEN.. rules
Networked schemes • Semantic nets • Conceptual graphs	**Structured schemes** • Scripts • Frames

(*i*) ***Logical schemes*** represent knowledge, using mathematical or orthographic symbols, inference rules and are based on precisely defined syntax and semantics.

(*ii*) In ***procedural schemes*** knowledge is represented as a set of instructions for problem-solving. That allows to modify a knowledge base easily and to separate a knowledge base from an inference mechanism.

(*iii*) ***Networked schemes*** use a graph to represent knowledge. Nodes of a graph display objects or concepts in a domain, but arcs define relationships between objects, their attributes and values of attributes.

(*iv*) ***Structured schemes*** extend networked representation by displaying each node in a graph as a complex data structure.

5.8 PROPERTIES FOR KNOWLEDGE REPRESENTATION SYSTEMS

The following properties should be possessed by a knowledge representation system :

(*i*) ***Representational Adequacy.*** This is the ability to represent the required knowledge ; the representation should have a set of well-defined syntax and semantics.

(*ii*) ***Inferential Adequacy.*** This is the ability to manipulate the knowledge represented to produce new knowledge corresponding to that inferred from the original.

(*iii*) ***Inferential Efficiency.*** This is the ability to direct the inferential mechanisms into the most productive directions by storing appropriate guides.

(*iv*) ***Acquisitional Efficiency.*** The ability to acquire new knowledge using automatic methods wherever possible rather than reliance on human intervention.

5.9 PARAMETERS OF A GOOD KNOWLEDGE REPRESENTING SCHEME

Important parameters of a good knowledge representing scheme are :

Ease of Representation

The ease with which a problem can be solved depends upon knowledge representation method.

Granularity of Representation

Granularity of knowledge representation can affect its usefulness, that is, how detailed the knowledge need to be represented. This will depend on the application and the use to which the knowledge will be put. *For example,* if a knowledge base about family a relationship is to be built and we start with 'cousin'. We may represent the definition of the relation as: your cousin is a child of sibling of you parent. For a female cousin your cousin is a daughter of a sibling of your parent and for a male cousin your cousin is a son of a sibling of your parent.

Expressiveness

An expressive representation scheme will be able to handle different types and levels of granularity of knowledge and knowledge structures and the relationships between them. Expressiveness also relates to the clarity of the representation scheme. Ideally the scheme should use a notation which is natural and usable both by the ***knowledge engineer*** and the ***domain expert.***

Explicitness

A good knowledge representation scheme be able to provide an explanation of its inference and allow justifications of its reasoning. The chain of reasoning should be explicit.

Efficiency

The scheme should not only support inference of new knowledge from old but must do so efficiently in order for new knowledge to be used. In addition, scheme should make possible efficient knowledge gathering and representation.

Effectiveness

Another measure of a good representation scheme is its effectiveness. In order to be effective the scheme must provide a means of inferring new knowledge from old. It should also be agreeable to computation, allowing adequate tool support.

Property Inheritances

World can be represented by facts and facts are related to each other. A fact may be specific instance of another more general fact. For example, "Black dogs bark" is a specific instance of the fact "All dogs bark". This is a case of property inheritance, in which properties of attributes of the main class are inherited by instance of that class. So we represent the knowledge that "dogs bark" and that "black dog is the dog", allowing us to deduce by inheritance the fact that "Black dogs bark".

5.10 ISSUES IN KNOWLEDGE REPRESENTATION

Below are listed issues that should be raised when using a knowledge representation technique :

Important Attributes

- Are there any attributes that occur in many different types of problem ?

There are two *instance* and *isa* and each is important because each supports property inheritance.

Relationships Among Attributes

- What about the relationship between the attributes of an object, such as, inverses, existence, techniques for reasoning about values and single valued attributes. We can consider an example of an inverse in

```
band(John Zorn, Naked City)
```

This can be treated as John Zorn plays in the band *Naked City* or John Zorn's band is *Naked City*.

Another representation is *band = Naked City*

```
band-members = John Zorn, Bill Frissell, Fred Frith, Joey Barron, ...
```

Choosing Granularity

At what level should the knowledge be represented and what are the primitives. Choosing the Granularity of Representation Primitives are fundamental concepts such as holding, seeing, playing and as English is a very rich language with over half a million words it is clear we will find difficulty in deciding upon which words to choose as our primitives in a series of situations.

If *Tom feeds a dog* then it could become :

```
feeds(tom, dog)
```

If *Tom gives the dog a bone* like :

```
gives(tom, dog, bone)
```

Are these the same ?

In any sense does giving an object food constitute feeding ?

```
If give(x, food) feed(x) then we are making progress.
```

But we need to add certain inferential rules.

In the famous program on relationships *Louise is Bill's cousin* How do we represent this? *louise = daughter* (*brother or sister* (*father or mother*(*bill*))) Suppose it is *Chris* then we do not know if it is *Chris* as a male or female and then son applies as well.

Clearly the separate levels of understanding require different levels of primitives and these need many rules to link together apparently similar primitives.

Obviously there is a potential storage problem and the underlying question must be what level of comprehension is needed.

Set of Objects

- How should sets of objects be represented ?

Finding Right Structure

- Given a large amount of knowledge stored in a database, how can relevant parts be accessed when they are needed ?

5.11 APPROACHES TO KNOWLEDGE REPRESENTATION

There are many techniques for knowledge representation. The important types of these techniques are as follows :

1. Simple Relational Knowledge

The simplest way of storing facts is to use a relational method where each fact about a set of objects is set out systematically in columns. This representation gives little opportunity for inference, but it can be used as the knowledge basis for inference engines.

- Simple way to store facts.
- Each fact about a set of objects is set out systematically in columns.
- Little opportunity for inference.
- Knowledge basis for inference engines.

Musician	Style	Instrument	Age
Saurabh	Rock	Guitar	35
Suresh	Jazz	Saxophone	43
Vivek	Rock	Trumpet	40
Pradeep	Jazz	Guitar	35
Kapil	Rock	Trumpet	38

Figure 5.2 Simple relational knowledge.

We can ask things like :

- Who plays guitar ?
- Whose style is Rock ?
- Who plays Jazz/Trumpet etc. ?

This sort of representation is popular in database systems.

2. Inheritable Knowledge

Relational knowledge is made up of objects consisting of

- Attributes
- Corresponding associated values.

One of the most useful forms of inference is property inheritance. In Property inheritance

- Elements inherit values from being members of a class.
- Data must be organized into a hierarchy of classes.

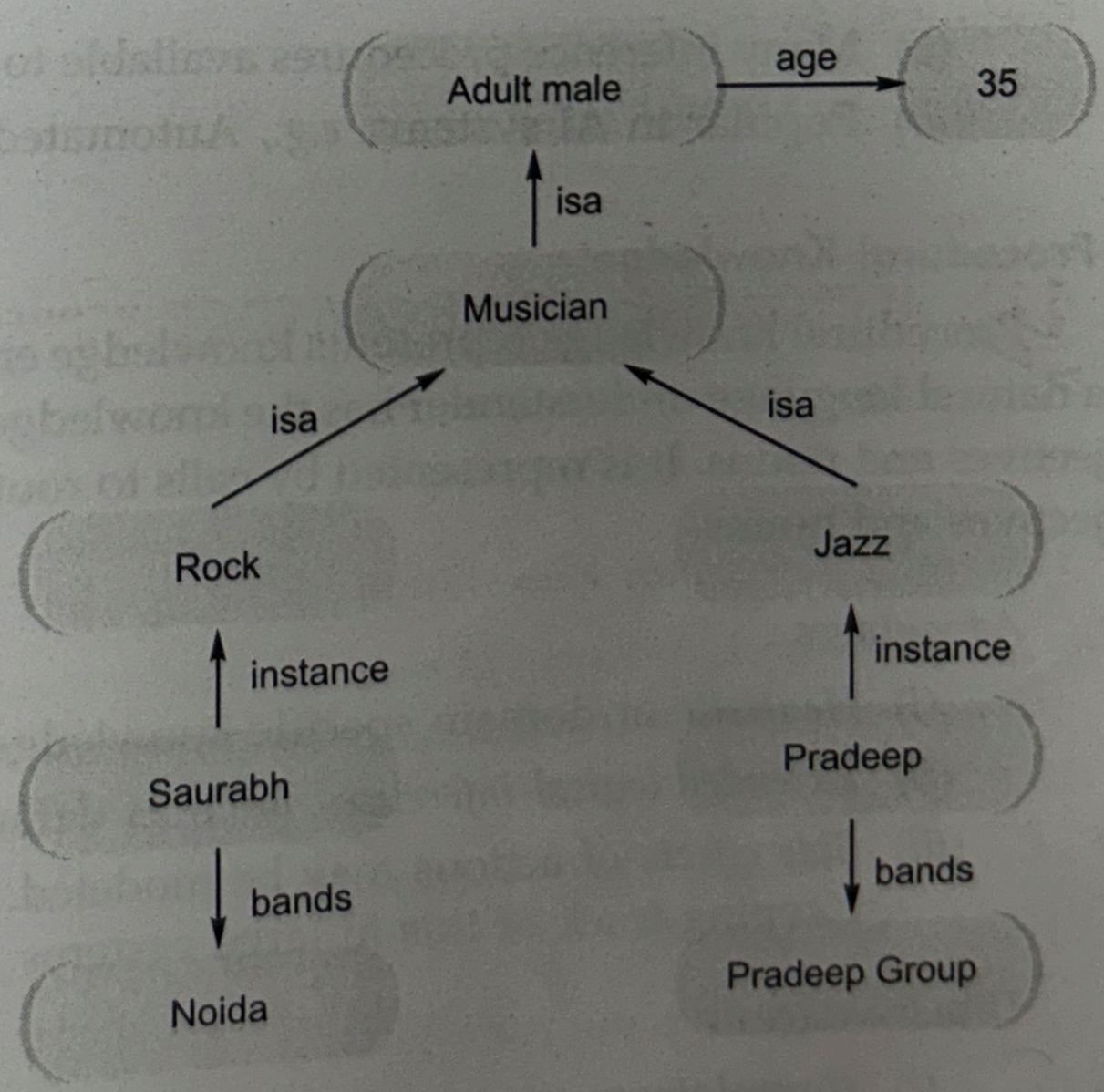

Figure 5.3 Inheritable knowledge.

- Boxed nodes — objects and values of attributes of objects.
- Values can be objects with attributes and so on.
- Arrows — point from object to its value.
- This structure is known as a **slot and filler structure**, semantic network or a collection of frames.

ALGORITHM *Property inheritance*

The algorithm to retrieve a value for an attribute of an instance object :

1. Find the object in the knowledge base
2. If there is a value for the attribute report it
3. Otherwise look for a value of instance if none fail
4. Otherwise go to that node and find a value for the attribute and then report it
5. Otherwise search through using *isa* until a value is found for the attribute.

3. Inferential Knowledge

Knowledge is useless unless there is some inference process that can exploit it. The required inference process implements the standard logical rules of inference. Inferential knowledge represents knowledge as *formal logic*.

For example : All dogs have tails

$$\forall x: \text{dog}(x) \rightarrow \text{hasatail}(x)$$

Advantages :

(*i*) A set of strict rules.

- Can be used to derive more facts.
- Truths of new statements can be verified.
- Guaranteed correctness.

(*ii*) Many inference procedures available to in implement standard rules of logic.

(*iii*) Popular in AI systems. *e.g.*, Automated theorem proving.

4. Procedural Knowledge

Procedural knowledge represents knowledge encoded in some procedures *e.g.*, a parser in a natural language understander has the knowledge that a *noun phrase* may contain articles, adjectives and nouns. It is represented by calls to routines that know how to process articles, adjectives and nouns.

Advantages :

(*i*) *Heuristic* or domain specific knowledge can be represented.

(*ii*) *Extended logical inferences*, such as default reasoning facilitated.

(*iii*) *Side effects* of actions may be modeled. Some rules may become false in time. Keeping track of this in large systems may be tricky.

Disadvantages :

(*i*) ***Completeness*** — not all cases may be represented.

(*ii*) ***Consistency*** — not all deductions may be correct. *e.g.*, If we know that *Fred is a bird* we might deduce that *Fred can fly*. Later we might discover that *Fred is an emu*.

(*iii*) Modularity is sacrificed. Changes in knowledge base might have far-reaching effects.

(*iv*) Cumbersome control information.

Exercises

1. What do you mean by knowledge representation ?
2. What is the need of knowledge representation ?
3. What are the various issues in knowledge representation ?
4. Discuss different approaches to knowledge representation.
5. Describe the various knowledge representation schemes used in AI. Discuss one representation scheme in detail.
6. Explain Production rules for knowledge representation.
7. What are the desirable properties of a good Knowledge Representation Systems ?