ME 620: Fundamentals of Artificial Intelligence

Lecture 13: Knowledge Representation and Reasoning

Shyamanta M Hazarika
Biomimetic Robotics and Artificial Intelligence Lab
Mechanical Engineering and M F School of Data Sc. & AI
IIT Guwahati

Intelligent Behaviour: Human vs. Artificial

- **Human Intelligence**
  - Most complex and mysterious phenomenon
  - Striking aspect of intelligent behaviour is that it is clearly conditioned by knowledge.
  - Decisions for a wide range of activities are based on what we know (or belief).
- **Intelligent behaviour** through computational means;
  - Knowledge Representation and Reasoning is concerned with how an agent uses what it knows in deciding what to do.
    - Structures for representing the knowledge.
    - Computational processes for reasoning with those structures.

Knowledge

data

- primitive verifiable facts, of any representation.
- Data reflects current world, often voluminous frequently changing.

information

- interpreted data

knowledge

- relation among sets of data (information), that is very often used for further information deduction.
- Knowledge is (unlike data) general.
- Knowledge contains information about behaviour of abstract models of the world.

Symbolic Vs. Connectionist AI

Symbolic AI

- Believed in developing an intelligent system based on rules and knowledge and whose actions were interpretable.
- The traditional symbolic approach, introduced by Newell & Simon in 1976 describes AI as the development of models using symbolic manipulation.

Connectionist AI

- Strived to build a computational system inspired by the human brain.

Knowledge Representation Hypothesis

Knowledge representation is an essential problem of symbolic-based artificial intelligence.

Any mechanically embodied intelligent process will comprise of structural ingredients, that

a. will represent the propositional account of knowledge the overall process exhibits
b. independently of such a formal semantics will play formal and causal role in performing the behaviour that manifests the knowledge.

Knowledge Representation Hypothesis
Brian Smith (1982)

Symbol System Hypothesis

- The **physical symbol system hypothesis** is a position in the philosophy of AI.

> A physical symbol system has the necessary and sufficient means for general intelligent action.
>
> Allen Newell and Herbert A. Simon

  - human thinking is a kind of symbol manipulation (because a symbol system is necessary for intelligence)
  - machines can be intelligent (because a symbol system is sufficient for intelligence).

Symbol System Hypothesis

- Knowledge may be represented as **symbol structures**
  - Essentially, complex data structures representing bits of knowledge (objects, concepts, facts, rules, strategies).
    - E.g., "red" represents colour red.
    - "car1" represents my car.
    - red(car1) represents fact that my car is red.
- **Intelligent behaviour** can be achieved through manipulation of symbol structures.
  - How to represent knowledge?
  - How to implement the process of reasoning?

Representation

- **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, a drawing of a hamburger on a sign might stand for a less immediately visible fast food restaurant;
- The type of representor that we will be most concerned with here is the formal **symbol**
  - A character or group of characters taken from some predetermined alphabet.
  - The digit "7," for example, stands for the number 7, as does the group of letters "VII"

What is a Knowledge Representation?

- Role I: A KR is a Surrogate
  - Imperfect Surrogates mean incorrect inferences are inevitable
- Role II: A KR is a Set of Ontological Commitments
  - Commitment begins with the earliest choices
  - The commitments accumulate in layers
  - Reminder: A KR is not a data structure
- Role III: A KR is a Fragmentary Theory of Intelligent Reasoning
  - What is intelligent reasoning?
    - Intelligent reasoning: the logical view and the psychological view
  - Which inferences are sanctioned?
    - Form and content of the answers
  - Which inferences are recommended?
- Role IV: A KR is a Medium for Efficient Computation
- Role V: A KR is a Medium of Human Expression

R. Davis, H. Shrobe, and P. Szolovits. What is a Knowledge Representation? *AI Magazine*, 14(1):17-33, 1993

What is a Knowledge Representation?

- A knowledge representation is **fundamentally a surrogate**
  - a *substitute* for the thing itself
  - used to enable an entity to determine consequences by reasoning about the world.
- It is **a set of ontological commitments**,
  - i.e., an answer to the question:

    In what terms should I think about the world?

What is a Knowledge Representation?

- It is a fragmentary theory of intelligent reasoning, expressed in terms of three components:

1. the representation's fundamental conception of intelligent reasoning;
2. the set of inferences the representation sanctions;
3. the set of inferences it recommends.

What is a Knowledge Representation?

- It is a medium for pragmatically efficient computation, i.e., the computational environment in which reasoning 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.

What is a Knowledge Representation?

- A **subarea of Artificial Intelligence** concerned with understanding, designing, and implementing ways of representing information in computers so that programs (agents) can use this information
  - to derive information that is implied by it,
  - to converse with people in natural languages,
  - to decide what to do next
  - to plan future activities,
  - to solve problems in areas that normally require human expertise.

What is Reasoning?

- **Reasoning** is the **use of symbolic representations** of some statements in order to derive new ones.
  - While statements are abstract objects, their representations are concrete objects and can be easily manipulated.
- Knowledge representation schemes are useless without the ability to reason with them.
  - Knowledge Representation and Reasoning!

How KR & R Works

- Intelligence requires knowledge
- Computational models of intelligence require models of knowledge
- Use formalisms to write down knowledge
  - Expressive enough to capture human knowledge
  - Precise enough to be understood by machines
- Separate knowledge from computational mechanisms that process it
  - Important part of cognitive model is what the organism knows!

How can knowledge be represented ?

- **Symbolic methods**
  - Declarative Languages (Logic)
  - Imperative Languages (C, C++, Java, etc.)
  - Hybrid Languages (Prolog)
  - Rules
  - Frames
  - Semantic Networks
  - …
- **Non – symbolic methods**
  - Neural Networks
  - Genetic Algorithms

Knowledge representation languages

- Rather than use general C++/Java data structures, use special purpose formalisms.
- A KR language should allow one to:
  - represent adequately the knowledge one needs for the problem
  - do it in a clear, precise and "natural" way.
  - allow one to reason on that knowledge, drawing new conclusions.

Requirements for KR language

- A KR language should have the following characteristics
  - Representational Adequacy
  - Clear syntax/semantics
  - Inferential adequacy
  - Inferential efficiency
  - Naturalness

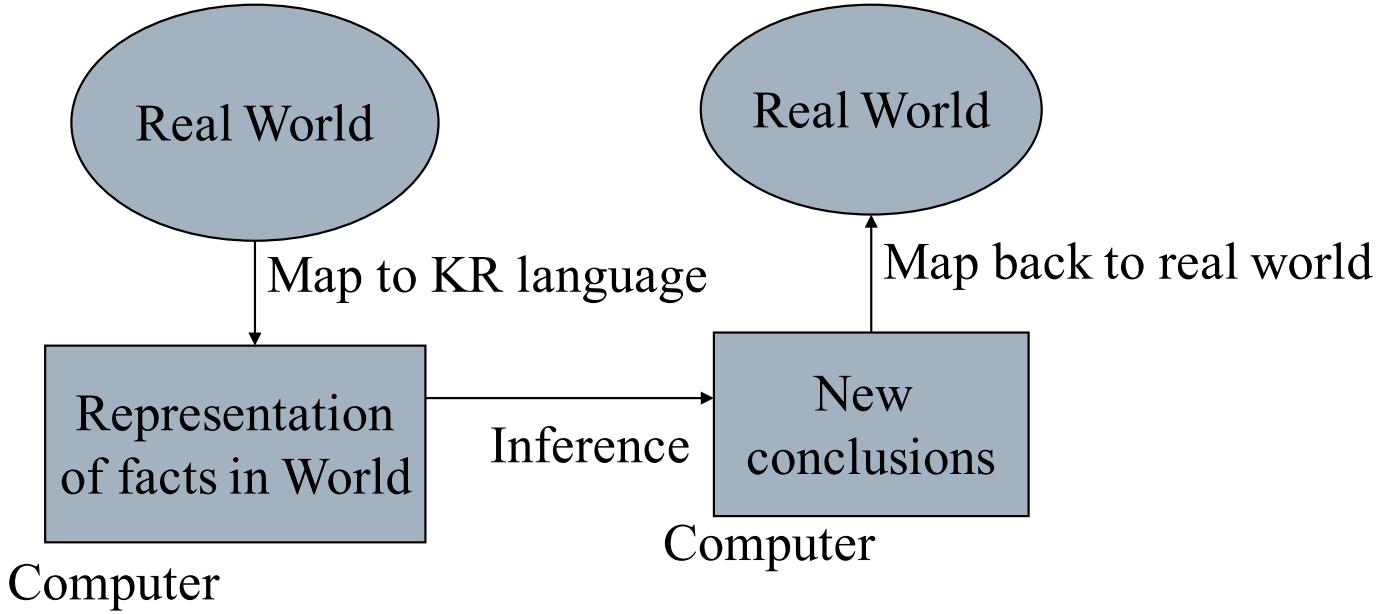
In practice no one language is perfect, and different languages are suitable for different problems.

Representational adequacy

- Consider the following facts:
  - Mohan believes no-one likes cricket.
  - Most children believe in Santa.
  - Mr. Kumar will have to finish his assignment before he can start working on his project.
- Can all be represented as a string?
  - May be Yes!
  - But hard then to manipulate and draw conclusions.
- **How do we represent these formally** in a way that can be manipulated in a computer program?
  - Some notations/languages only allow you to represent certain things.
  - Time, beliefs, uncertainty, all hard to represent.

Well-defined syntax/semantics

- Knowledge representation languages should have **precise syntax** and **semantics**.
- You must know exactly what an expression means in terms of objects in the real world.



Syntax of a KR language

- Need to specify which **groups of symbols**, **arranged in what way**, are to be considered **properly formed**.
  - In English, for example, the string of words

    The cat my mother loves is a well-formed phrase
    The my loves mother cat is not a well-formed phrase.
- The **syntax** consists of a set of symbols used by the language and a set of rules according to which the symbols can be combined to form proper sentences.

Semantics of a KR language

- For a KR language, need to specify what the **well-formed expressions** are supposed to mean.
- The **semantics** determine a mapping between symbols, combinations of symbols, propositions of the language and concepts of the world to which they refer
- A proposition in a KR language does not mean anything on its own
  - The semantics (i.e. the meaning) of the proposition must be defined by the language author through an **interpretation**

Inferential Adequacy

- Representing knowledge not very interesting unless you can **use it to make inferences**:
  - Draw new conclusions from existing facts.
    - "If its raining John never goes out" + "It's raining today" so..
  - Come up with solutions to complex problems, using the represented knowledge.
- Inferential adequacy refers to **how easy it is to draw inferences** using represented knowledge.
- Representing everything as natural language strings has good representational adequacy and naturalness, but very poor inferential adequacy.

Inferential Efficiency

- You may be able, in principle, to **make complex deductions** given knowledge represented in a sophisticated language.
  - But it may be just too inefficient.
  - Generally the more complex the possible deductions, the less efficient will be the reasoner.
- Need representation and inference system sufficient for the task, without being hopelessly inefficient.

Natural representation scheme

- Also helpful if our representation scheme is quite **intuitive** and **natural** for human readers!
- Could represent the fact that my car is red using the notation:
  - "`xyzzy ! Zing`"
  - where `xyzzy` refers to redness, `Zing` refers to by car, and ! used in some way to assign properties.
- But this wouldn't be very helpful; compared with
  - car(smh) AND colour(red).
  - where car(x) refers to owner x; colour(x) refer to `x' colour.

Desired Features of KR languages

- **Epistemological Level**
  - Clarity
  - Expressiveness
- **Logical Level**
  - Elegant syntax & semantics
  - Decidability / Tractability
  - Sound and complete inference mechanism
- **Implementation Level**
  - Space & Time efficiency
  - Extensibility