

# Knowledge Representation

- If we have specific knowledge about a problem it is easier to solve it
- We can combine this domain knowledge with the general knowledge that we have about problem solving
- This knowledge allows to the different AI algorithms to guide the search and reduce the cost to obtain a solution
- Problems
  - How to choose a representation formalism that allows an easy translation from the real world to the represented world?
  - How is this representation used efficiently?

# Differences between information and knowledge (1)

- **Information** is defined as a set of basic facts, without interpretation, obtained as the definition of the problem
  - For example:
    - The numerical data from a blood test
    - The data from the sensors of a chemical process
- **Knowledge** is defined as a set of high level facts that model structuredly the experience from a domain or that are obtained from the interpretation of the basic facts
  - For example:
    - The interpretation of the values of a blood test or of the sensors from a chemical process, saying if they are normal, high, low, problematic, dangerous, ...
    - The data structures and methods to diagnose patients from the interpretation of a blood test or to aid to the decision process in the management of a chemical plant

## Differences between information and knowledge (2)

- AI programs need different kinds of knowledge that are not available from databases and other information sources:
  - Knowledge about the objects of a domain and their relationships
  - Knowledge about the processes that involve these objects or that are useful to manipulate them
  - Knowledge that is difficult to represent as basic facts as intensionality, causality, goals, temporal information, “common sense” knowledge, ...
- Intuitively

*Knowledge = Information + Interpretation*

# Knowledge representation

- To represent knowledge we need
  - Its structure
  - What for the intelligent agents will use it
  - How it will be used by intelligent agents
  - How it will be acquired
  - How it will be stored and manipulated
- Unfortunately there are not complete answers to these questions from a biological or neurophysiological point of view
  - We will see formalisms that simulate the acquisition, structure and manipulation of knowledge that will allow us to build intelligent agents

# Knowledge Representation formalisms

- A **Representation formalism** is an mechanism to represent the real world in a computer
- It is important to keep in mind the difference between
  - The real world (what we want to represent) → **Domain**
  - Its representation → **many representation formalism**
- A representation formalism can be seen as a combination of:
  - Data structures to code the problem that an intelligent agent is solving → **Static part**
  - Data structures that store the domain knowledge of the problem and actions to manipulate knowledge consistently with its interpretation → **Dynamic part**

# Knowledge Representation formalisms: Static part

- The static structures include
  - The data structure that represents the knowledge of the problem
  - The actions that allow to create, modify or delete elements from the data structure
  - Predicates that allow to query the data structure
  - Semantics of the data structure: a semantic correspondence between the real world and the representation is chosen

*R(Elements from representation, Real world)*

# Knowledge Representation formalisms: Dynamic part

- The dynamic part is composed of:
  - Data structures that store the domain knowledge
  - Actions that allow
    - To interpret the data from the problem (the static part) using the domain knowledge (the dynamic part)
    - Control the use of the data: Control strategies
    - Acquire new knowledge

# Incompleteness of knowledge representation

- Any representation will be always incomplete because:
  - **Change:** The real world is dynamic and changes every moment, but our representation only captures an instant
  - **Volume:** lots (too much) of knowledge to represent → partial view
  - **Complexity:** Reality has lots of details
- The problem of change is related to the processes of acquiring and maintaining the knowledge in the representation (Frame Problem)
- The problems of volume and complexity are related with the granularity of the representation

# Properties of Knowledge Representation formalisms

A Knowledge Representation formalism has the following properties

- Of the formalism itself
  - **Representational Adequacy:** It must support representation of all kinds of knowledge required by the problem
  - **Inferential Adequacy:** It must allow to manipulate the knowledge represented to obtain new knowledge from the current one

# Properties of Knowledge Representation formalisms

- Related to the use of the representation
  - **Inferential Efficiency:** It needs to be able to infer new things from the knowledge easily. The ability to incorporate additional information that can be used to focus the attention of the inference mechanisms to promising areas (metaknowledge)
  - **Acquisitional Efficiency:** It must allow to incorporate new knowledge to the representation easily. Ideally an intelligent agent should be able to obtain new information autonomously and incorporate it to the representation

# Typology of knowledge

- Declarative knowledge
  - The knowledge is represented independently from the mechanisms that will use it
  - The control of the adequate use of the knowledge is performed
    - by means of general purpose heuristics that determine the best way to use the knowledge
    - by means of the use of knowledge about the control of the use of the knowledge that guides the solving mechanism
  - Kinds of declarative knowledge
    - Relational Knowledge
    - Inheritable Knowledge
    - Inferential Knowledge
- Procedural Knowledge
  - The knowledge explicitly defines how to use it

# Simple relational knowledge

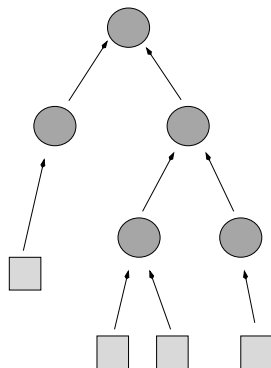
- The easier way to represent declarative facts is to use relations that can be represented as tables (as in databases)
  - Ex: Information about the clients of a company

Client	Address	Sales	...
A. Perez	Av. Diagonal	5643832	
J. Lopez	c/ Industria	430955	
...			

- Problem: There is not much knowledge represented
- We need some mechanisms to improve the representation → Inference engine: obtains new knowledge from the representation
  - Ex: mean sales of a product, best clients, client typology
- Databases can be useful for Knowledge Based Systems.

# Inheritable Knowledge

- It is useful to structure the knowledge hierarchically (taxonomical hierarchy)
- The goal is to represent the knowledge as a graph or a tree and to use generalization and specialization as representation goal
  - The nodes are concepts/classes, the edges are relations
    - is-a: class-class relation
    - Instance-of: class-instance relation
- The reasoning mechanism is the inheritance of properties and values
  - Simple/multiple inheritance
  - Default values



# Inferential knowledge

- The knowledge is described using logic
- The semantics of the logic connectives and inference mechanisms (for example Modus Ponens rule) can be used to obtain new knowledge

$$\forall x, y : \textit{person}(x) \wedge \neg \textit{underage}(x) \wedge \neg \textit{job}(x, y) \rightarrow \textit{unemployed}(x)$$

- The inference mechanisms for first order logic are the different algorithms from automatic theorem proving

# Procedural knowledge

- This knowledge includes control information in its representation and/or specific procedures needed to use it
  - Programs: The knowledge is defined as algorithms that allow to infer new knowledge from facts
    - Ex: `Date_of_birth= DD-MM-YYYY; function Age (Date_of_birth: date)`
  - Production rules: if some condition holds some actions/inferences are performed
    - Ex: IF conditions THEN actions
- This kind of knowledge is more computationally efficient, but is more difficult to obtain new knowledge from it and it is complex to acquire/modify