

Sistemi Basati sulla Conoscenza 2011 / 2012

Umberto Straccia

ISTI-CNR, Pisa, Italy
<http://www.straccia.info>
straccia@isti.cnr.it

What is Knowledge Representation?

Basics on Classical Logical Languages

- Propositional Logic
- First-Order Logic

Knowledge Engineering Pitfalls

Basics on Semantic Web Languages

- Web Ontology Languages
- RDF/RDFS
- Description Logics
- Logic Programs

Uncertainty and Vagueness Basics

- Sources of Uncertainty and Vagueness
- Uncertainty vs. Vagueness: a clarification
- Uncertainty & Vagueness: Basic Concepts
- Uncertainty & Logic
- Probability & Logic
- Vagueness & Logic

Fuzzyness in Semantic Web Languages

- The case of RDFS
- The case of Description Logics
- The case of Logic Programs

What is Knowledge Representation?

What is Knowledge?

Knowledge is defined (Oxford English Dictionary) variously as:

- ▶ facts, information, and skills acquired by a person through experience or education;
- ▶ the theoretical or practical understanding of a subject;
- ▶ what is known in a particular field or in total –facts and information;
- ▶ awareness or familiarity gained by experience of a fact or situation.

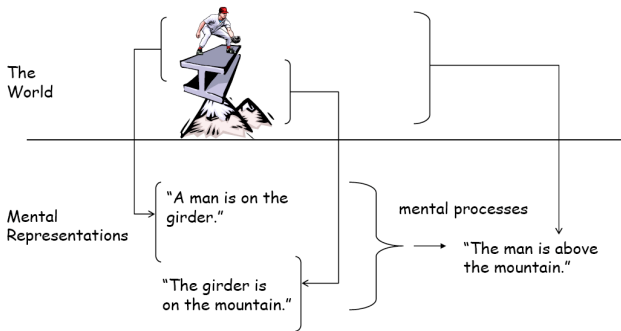
Source: <http://en.wikipedia.org/wiki/Knowledge>

However,

- ▶ philosophical debates in general start with Plato's formulation of knowledge as "justified true belief";
 - ▶ in order to count as knowledge, a statement must be justified, true, and believed;
 - ▶ **belief** is the psychological state in which an individual is convinced of the truth or validity of a statement without the ability to adequately prove their main contention for other people who may disagree.
- ▶ there is no single agreed definition of knowledge presently, nor any prospect of one, and there remain numerous competing theories.

As a matter of fact,

- ▶ the "manipulation of knowledge" seems to be at the heart of what we as intelligent beings do



Knowledge Representation and Reasoning

- ▶ To try to model the **manipulation of knowledge** in a computer we need to:
 - ▶ **represent** knowledge using **symbol** structures; and
 - ▶ perform **formalized** versions of **reasoning**.
- ▶ This means that we need carefully specified languages for the representation of knowledge.
- ▶ **Knowledge Representation and Reasoning** deals with the problem to represent and reason with knowledge.

What is a Symbol?

From Merriam-Webster's Collegiate:

- ▶ Something that stands for or suggests something else.
- ▶ An arbitrary conventional sign used in writing or printing to represent:
 - ▶ operations
 - ▶ quantities
 - ▶ elements
 - ▶ relations
 - ▶ qualities

Symbols standing for things in the world

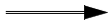


first aid



women

"John"



John

"John loves Mary"

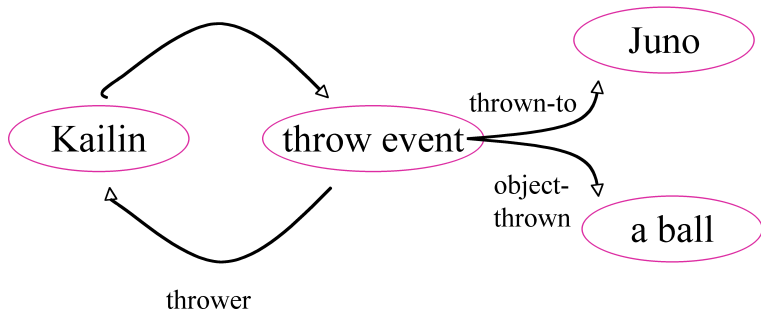


the proposition that
John loves Mary

What is a Representation?

- ▶ What does it mean to represent something?
- ▶ Identify the following:
 - ▶ symbols
 - ▶ context
 - ▶ observer
- ▶ Representation is the association of symbols with conceptual objects or ideas in a given context
- ▶ The observer sets up a correspondence between the symbols and the meanings
- ▶ The concept associated to a symbol may mean different things in different contexts
 - ▶ The question “Is there any salt?”
 - ▶ Asked of a waiter at a restaurant
 - ▶ Asked of an environmental scientist at work

Kailin threw the ball to Juno
did-action



Semantics: The Meaning of Symbols

- ▶ Semantics: assigning meanings to symbols and expressions
- ▶ Usually involves defining
 - ▶ objects
 - ▶ properties of objects
 - ▶ relations between objects
- ▶ More detailed versions include (among others)
 - ▶ events
 - ▶ time
 - ▶ places
 - ▶ measurements (quantities)

General Properties of Good Representations

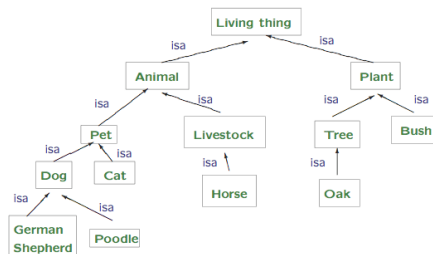
- ▶ They make the **important objects and relations explicit**
 - ▶ you can see what's going on at a glance
- ▶ They expose **natural constraints**
 - ▶ you can express the way one object or relation influences another
- ▶ They **bring objects and relations together**
 - ▶ you can see all you need to see at one time
- ▶ They **suppress irrelevant detail**
 - ▶ you can keep rarely used details out of sight, but still get to them when necessary

- ▶ They are **transparent**
 - ▶ you can understand what is being said
- ▶ They are **complete**
 - ▶ you can say all that needs to be said
- ▶ They are **concise**
 - ▶ You can say what you need to say efficiently
- ▶ They are **fast**
 - ▶ you can store and retrieve information rapidly
- ▶ They are **computable**
 - ▶ you can create them with an existing procedure

Fundamental Properties of Representation

- ▶ A **lexical** part that determines which symbols are allowed in the representation's vocabulary
- ▶ A **structural** part that describes constraints on how the symbols can be arranged
- ▶ A **procedural** part that specifies access procedures that enable you to create descriptions, modify them, and to answer questions using them
- ▶ A **semantic** part that establishes a way of associating meaning with the description

Example: ISA-Hierarchy



- ▶ **Lexical** part:
 - ▶ nodes, denoting objects (mathematically, sets)
 - ▶ links, denoting (binary) relations between objects
 - ▶ link labels, denote particular relations (mathematically, subset relation, \subseteq)
- ▶ **Structural** part: directed graph, with labeled links
- ▶ **Procedural** part:
 - ▶ constructors to make,
 - ▶ readers to answer questions,
 - ▶ writers to alter, and
 - ▶ erasers to delete nodes and links
- ▶ **Semantic** part: meaning of nodes (sets) and links (subset relation)

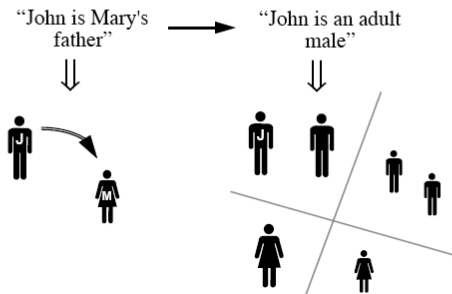
What is Reasoning?

- ▶ Representation alone is not useful
- ▶ We want to be able to access represented knowledge and to **process** it
 - ▶ access alone is, in general, insufficient
 - ▶ **implicit** knowledge has to be made explicit
- ▶ **Deduction methods**
 - ▶ the results should only depend on the semantics . . .
 - ▶ and not on accidental syntactic differences in representations

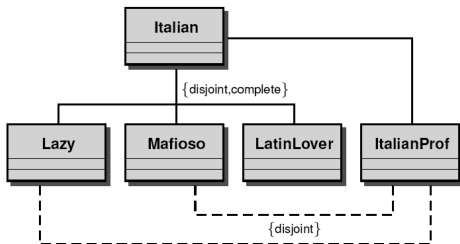
What is Reasoning?

Manipulation of symbols encoding propositions to produce representations of new propositions

Analogy: arithmetic "1011" + "10" → "1101"
 ↓ ↓ ↓
 eleven two thirteen

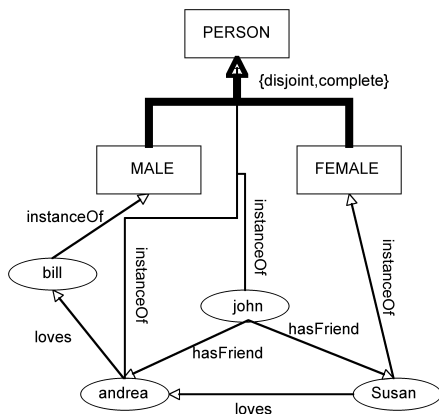


Example: reasoning at the conceptual level



- ▶ It is true that
“an ItalianProf is a LatinLover”?
- ▶ Yes, argument: an *ItalianProf* is Italian, hence has to be either *Lazy*, *Mafioso* or *Latinlover*. But, it cannot be the former two, so it has to be last one, i.e. a *LatinLover*.

Example: reasoning at the individual level



- ▶ Is it true that
“John has a female friend loving a male person”?
- ▶ Yes, argument: because andrea is either male or female

Requirements for a KR language

- ▶ First, we need **representational adequacy**
 - ▶ Can I represent the pieces of knowledge I need to?
 - ▶ Can I represent the pieces of knowledge I need to in such a way that reasoning can be automated?
- ▶ English is highly expressive in representing knowledge:
 - ▶ “Ophelia believes that all sensible people dislike eating pies”
- ▶ However automating reasoning based on English language representations is just about impossible at present.
 - ▶ How would we write a program that takes this statement and when told
 1. “Marta is really sensible”
 2. “Marta is a funny sort of person”infers that
 1. “Ophelia believes Marta dislikes eating pies”?

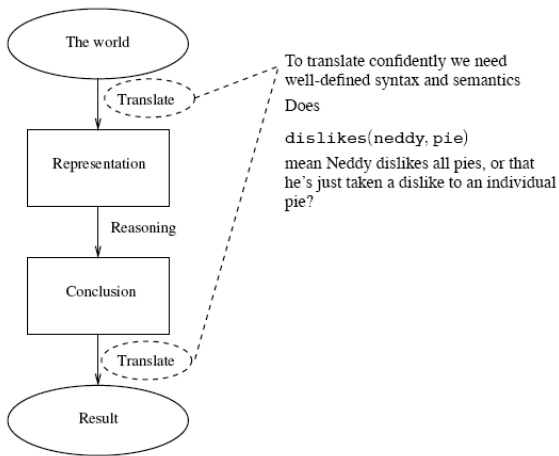
On the other hand:

$$\begin{array}{c} \text{person(marta)} \\ \text{sensible(marta)} \end{array} \quad \forall x \text{ sensible}(x) \wedge \text{person}(x) \implies (\forall y \text{ pie}(y) \implies \text{dislikes}(x, y))$$

is something for which reasoning can be automated.

In addition to needing an expressive language, the language needs to be clearly defined:

- ▶ We need syntax and semantics:
 - ▶ **Syntax**: defining when a statement in the language is well-formed;
 - ▶ **Semantics**: specifying what a statement in the language means.
- ▶ English is again not good here from the point of view of automation.
- ▶ If possible, we also want the representation to be natural in the sense that it is reasonably easy to understand and deal with.

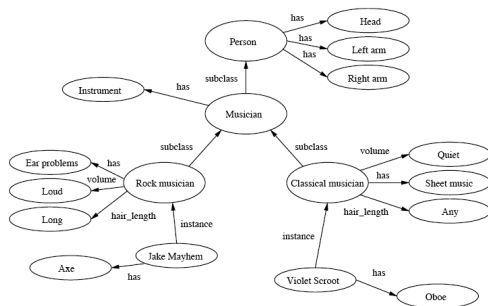


Inferential adequacy and inferential efficiency

- ▶ It is not always possible, and it's certainly not desirable, to store all knowledge as explicit facts
 - ▶ Knowing that “all dogs smell bad” should allow us to infer that “fido smells bad”
 - ▶ We don't want to store a piece of knowledge for every possible dog
- ▶ However, more complex inferences are likely to take longer
- ▶ So as usual, there is a trade-off.

KR Languages

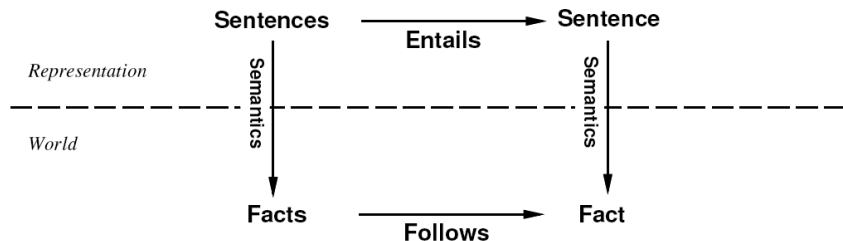
- ▶ Various candidates have been proposed for KR languages
 - ▶ Semantic Networks, Frame-based systems, Conceptual Graphs, etc.
- ▶ Example of Semantic Network



- ▶ There is **no single most** adequate knowledge representation formalism for everything
- ▶ Main points for selecting a representation formalism: **what** should be represented, **how** should the knowledge be **processed**

- ▶ One set of proposals is that formal **logic** be used as a basic framework for such languages
- ▶ Logic consists of
 - ▶ **Syntax**: The atomic symbols of the logical language, and the rules for constructing well-formed, non atomic expressions (sentences) of the logic
 - ▶ **Semantics**: The meanings of the atomic symbols of the logic, and the rules for determining the meanings of non-atomic expressions of the logic
 - ▶ **Proof theory**: determines which sentences are valid inferences from other sentences

Logic



- ▶ A logic allows the axiomatization of the domain information, and the drawing of conclusions from that information
 - ▶ Syntax
 - ▶ Semantics
 - ▶ Logical inference = reasoning

Logics and KR

- ▶ There are many logics for KR
 - ▶ differing mostly in terms of how much they can express
- ▶ Using logic seems reasonable in that the language was designed not to have the “flaws” of natural languages
 - ▶ i.e., we can engineer our way around ambiguity, etc.
- ▶ Using logic may poses problems as
 - ▶ designing a efficient valid inference procedure may not possible for all but the simplest logics
- ▶ Note: if we agree on a logic, that is just the beginning:
 - ▶ we still have to figure out what to represent

Important Questions

- ▶ **Expressive Power** of the logic:
 - ▶ able to represent the problem
- ▶ **Correctness** of inference procedure:
 - ▶ no false conclusions are drawn
- ▶ **Completeness** of inference procedure:
 - ▶ all correct conclusions are drawn
- ▶ **Decidability** of inference problem:
 - ▶ there exists a (terminating) algorithm to compute the inference
- ▶ **Complexity**:
 - ▶ computing resources needed for computing the solution

Summary: What is a Logic

Clearly distinguish the definitions of:

- ▶ the **formal language** :
 - ▶ Syntax
 - ▶ Semantics
 - ▶ Expressive Power
- ▶ the **reasoning problem**:
 - ▶ Decidability
 - ▶ Computational Complexity
- ▶ the **problem solving procedure**:
 - ▶ Soundness and Completeness