KDI : Knowledge and Data Integration

Stratified Representation

Fausto Giunchiglia, Mayukh Bagchi

Contents

Introduction

Motivation

To represent the *world*, we use:

- Words
- Sounds
- Pictures
- .....

Representation = Concepts + Language + Knowledge + Data

For example:

Starry Night is an oil on canvas painting by Vincent van Gogh.

Contents

What are concepts?

Concepts are *alinguistic units* of human *perception* and understanding generalized from particular instances of substances which we observe.

They allow us to give an account of phenomena such as *knowledge acquisition and representation*, language understanding, inference, and categorization.

Examples:- Car, my car with nameplate FP372MK.

Contents

What is language?

Language is a *tool for representing* what is the case in the world.

We take language as a *set of terms*, each term being associated a meaning, standing for what is the case in the world.

Examples:- English, Italian, Bengali, Hindi

Contents

What is knowledge?

Knowledge is a *tool for stating* what is the case in the world.

We take Knowledge as a set of *entity types (etypes)*, each associated with a set of *properties* (Object Properties and Data Properties).

Examples:-

In the automobile domain,

etype = {Car, Motorcycle}; properties = {Body Style, Model Date, VIN}

Contents

What is data?

Data is knowledge about *individuals*.

We take data as *a set of entities*, each of a given etype, each associated with a set of *property values*.

Examples:-

My car with *VIN: FP372MK* having body style *Coupé* and model date *2020-11-25*

Contents

The problem of representation diversity

Stratification of Diversity

We take *Representation Diversity* to mean semantic heterogeneity, as organized in the *four layers*:

Conceptual Diversity (L1)

Language Diversity (L2)

Knowledge Diversity (L4)

Data Diversity (L5)

Car				
Nameplate	schema: speed	schema: fuelCapacity	schema: fuelType	schema: modelDate
FP372MK	150	62	Petrol	2020-11-25

Vettura		
Targa	Velocità	Tipo di corpo
FP372MK	158	Coupé

Vehicle			
vso:VIN	vso:feature	vso:modelDate	vso:speed
FP372MK	Armrest	2020-11-25	155.0

Summary

- We saw the different components which make up any representation
- We introduced the precise role of concepts, language, knowledge and data in terms of knowledge representation
- We saw how the above four components generate different representations, even of the same intended meaning (but not only), in different datasets, thus generating the problems of

  KNOWLEDGE AND DATA INTEGRATION

  which is what we are going to learn how to manage in this course

KDI : Knowledge and Data Integration

- Fausto Giunchiglia, Mayukh Bagchi
- Stratified Representation