ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Upper ontologies and other things

Prof. Ing. Federico Chesani

DISI

Department of Informatics – Science and Engineering

Motivation

- There are many different types of knowledge
- However, there is a limited number of very high-level concepts that appear constantly

- A step further: categories and objects as the most intuitive type of knowledge we (humans) exploit
  - which characteristics?
  - which properties?

Goal of this lesson

- A limited number of very high-level concepts that appear constantly -> Upper ontologies
- A step further: categories and objects as the most intuitive type of knowledge we (humans) exploit
  - Objects and categories
  - reification, for representing them in FOL
  - disjunctness, exhaustive decomposition, partition
  - physical composition (composed objects vs. bunch of)
  - measures
  - things vs. stuff, intrinsic vs. extrinsic properties

Disclaimer and Further reading

Reading:

- AIMA, chapter 10, Sections 10.1, 10.2, 10.3

Further reading:

Upper Ontologies

What is an Ontology?

An ontology is a formal, explicit description of a domain of interest.

- formal: it should be described using a language with clear and non-ambiguous semantics
- explicit: the information should be readily available or, in the worst case, derivable in a finite time with a sound procedure
- description: it should provide us information… more interesting, the information it provides define the goal of the ontology itself
- domain: such description should be related to some topic of interest, and in turn it should be related to the ontology goal.

Ontologies… do we use them?

- We (humans) continuously make use of ontologies
- It seems that is an evolutionary trait of our brain, and an economic way for
  - storing information
  - communicating information

Generally speaking, we (humans) tend to organize our knowledge through two notions:

- Categories
- Objects (belonging to the categories)

Example: university employers

Problem: the new university director wants to organize all the employees, with the only purpose of the salary management.

University employees belong to:

- the administratives group;
- the technicians group;
- the researchers/teachers group.

- administrative and technical staff are organized into four different groups (B, C, D, EP, each with a different salary)
- researchers are organized in many different groups (because we are in Italy); roughly speaking: researchers, associates and full professors.

Example: university employers

Few questions:

- Does it make sense to keep the distinction between administrative and technical staff?
  - if the focus is the salary only… may be no

- Is it possible that a person belongs at the same time to two categories?

- Each university employee has a fiscal code (for taxation).
- To be precise, every resident in Italy has a fiscal code.
- This could suggest that there is a further category, the one of "resident in Italy"
- In turn, there could be a more general category of "Person"
- What is a "Person"? **WRONG QUESTION**: we started from the salary, let's focus on the salary…

Example: travel means

Citizen of Bologna use several different travel means. The major wants to put some order and evaluate some new green policy.

- Travel means can be divided into: public vs. private transports.

- Travel means can be divided into:
  - those that fly
  - those that sail
  - those that run on railways
  - those that run on wheels on concrete and asphalt

Example: travel means

- Travel means can be divided into: public vs. private transports.
- Travel means can be divided into:
  - those that fly
  - those that sail
  - those that run on railways
  - those that run on wheels on concrete and asphalt

- Apparently, there are two orthogonal dimensions
- Automobiles belongs to a category
- My car is an instance of one category

Categories, hierarchies, and objects

- There are categories (staff vs. professors)
- Categories can be more or less general (person vs. professor)
- Categories can be organized hierarchically

- Objects can be instances of categories
- Objects can belong to more categories at the same time

Upper ontologies

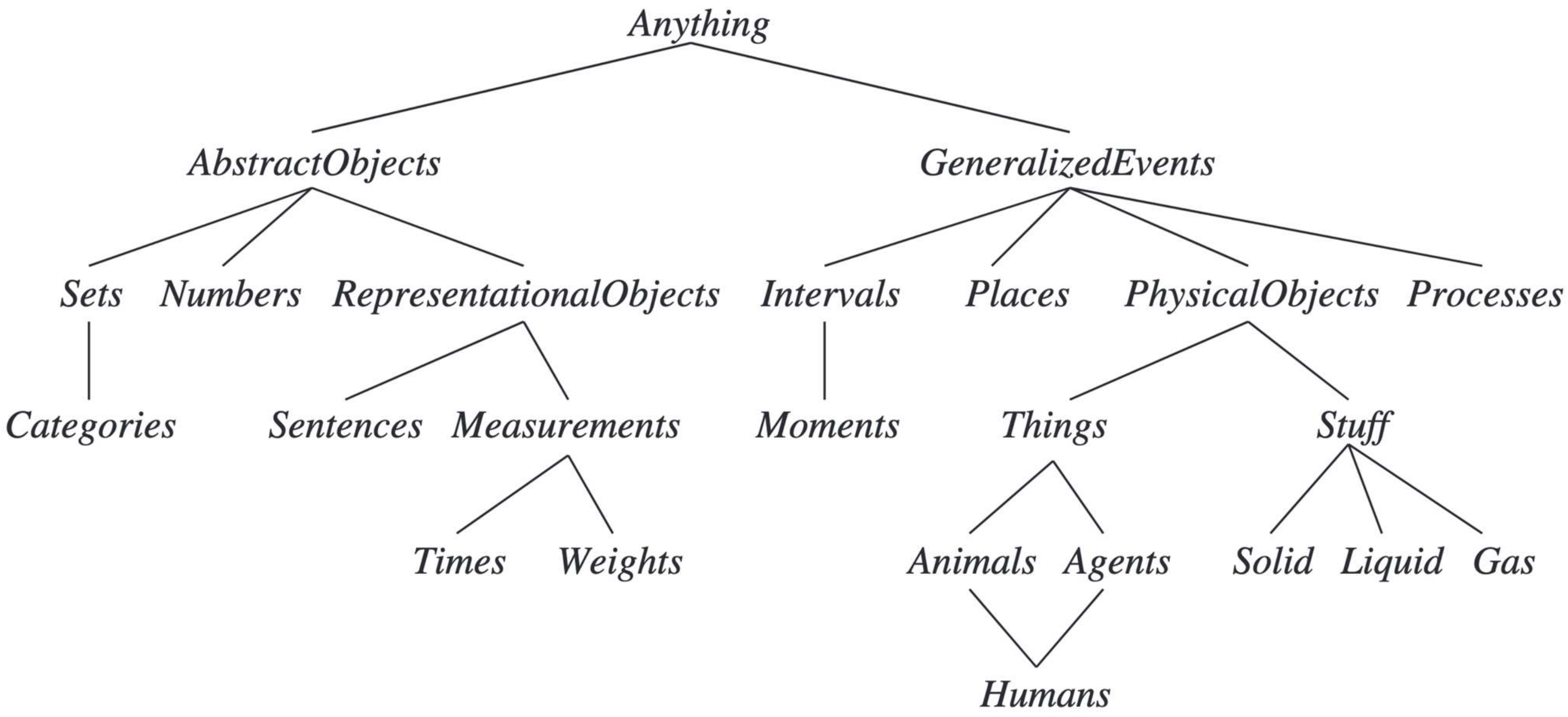
Among the categories, there can be:

- very general categories (e.g., Person)
- very specific ones (e.g., Employee of the University of Bologna in the administrative part, category D)

Generally speaking, this distinction between "very general" and "very specific" can be extended to the domain we want to describe, and then to the ontology that describes it.

An upper ontology is an ontology that focuses on the largest, more general, more high level, most general purpose domain.

Upper ontologies – an example



From AIMA, Chapter 10.1. The link between concepts must be interpreted as "specialization".

Upper ontologies

How special purpose ontologies relate with general purpose ontologies?

- Starting from special purpose, you can always move up (through *generalisation*), towards more general concepts
- "Do all these ontologies converge on a general-purpose ontology?" "Maybe" [cit. AIMA]

At least two characteristics of upper ontologies:

- they should be applicable to (almost) any special domain
- we should be able to combine different general concepts without incurring in inconsistencies (e.g., time and space concepts should work together...)

Upper ontologies

Roughly speaking, at least four different approaches to upper/general ontologies:

1) Ontologies created by philosophers/logician/AI researchers

Example: The Cyc ontology for Common sense reasoning

Douglas B. Lenat, R. V. Guha, Karen Pittman, Dexter Pratt, and Mary Shepherd. 1990. **Cyc: toward programs with common sense.** Commun. ACM 33, 8 (Aug. 1990), 30–49. https://doi.org/10.1145/79173.79176

Currently, commercial product: https://cyc.com/

Also some open version for academic purposes, namely OpenCyc

Upper ontologies

2) Ontologies created by automatically extract knowledge from already well-structured datasets

Example: DBpedia, extracted from Wikipedia, and explicitly using the structure of wikipedia entries.

Lehmann, Jens et al. '**DBpedia – A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia**'. 1 Jan. 2015 : 167 – 195.

Open-source, free ontology

Upper ontologies

3) Ontologies created by text documents

Example: TextRunner, extracted from text in web pages (NLP).

Alexander Yates, Michael Cafarella, Michele Banko, Oren Etzioni, Matthew Broadhead, and Stephen Soderland. 2007. **TextRunner: open information extraction on the web**. In Proceedings of Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations (NAACL-Demonstrations '07). Association for Computational Linguistics, USA, 25–26.

https://openie.allenai.org/

Upper ontologies

4) By using crowd-sharing information

Example: OpenMind

Push Singh, Thomas Lin, Erik T. Mueller, Grace Lim, Travell Perkins, and Wan Li Zhu. 2002. **Open Mind Common Sense: Knowledge Acquisition from the General Public**. In On the Move to Meaningful Internet Systems, 2002 - DOA/CoopIS/ODBASE 2002 Confederated International Conferences DOA, CoopIS and ODBASE 2002. Springer-Verlag, Berlin, Heidelberg, 1223–1237.

https://www.media.mit.edu/projects/open-mind-common-sense/overview/

Closed project.

Upper ontologies – a famous example

The foundational ontology DOLCE

Outcome of a successful EU-funded research project

http://www.loa.istc.cnr.it/dolce/overview.html

It is definitely a upper, foundational ontology…

2.4 Endurants and Perdurants

Classically, *endurants* (also called *continuants*) are characterized as entities that are 'in time', they are 'wholly' present (all their proper parts are present) at any time of their existence. On the other hand, *perdurants* (also called *occurrents*) are entities that 'happen in time', they extend in time by accumulating different 'temporal parts', so that, at any time t at which they exist, only their temporal parts at t are present.[12] For example, the book you are holding now can be considered an endurant because (now) it is wholly present, while "your reading of this book" is a perdurant because, your "reading" of the previous section is not present now. Note that it is possible to distinguish between 'ordinary objects' (like the book) and 'events or process' (like 'the reading of the book') even when the domain contains perdurants only. In this latter case, one relies on properties that lie outside spatio-temporal aspects. Indeed, one can assume that four-dimensional entities do not need to have different spatio-temporal locations. A person and its life (both taken to be 4D entities) share the same space-time region but differ on other properties since, for instance, color, race, beliefs and the like make sense for person only.

Endurants and perdurants can be characterized in a different way. Something is an endurant if (*i*) it exists at more than one moment and (*ii*) its parts can be determined only relatively to something else (for instance time)[49]. In other words, the distinction is based on the different nature of the parthood relation: endurants need a time-indexed parthood, while perdurants do not. Indeed, a statement like "this keyboard is part of my computer" is incomplete unless you specify a particular time, while "my youth is part of my life" does not require such a specification.[13]

Categories

Categories… do we need them?

- It is commonly accepted that humans represent the knowledge in terms of categories and objects belonging to

  However, this alone would not be a strong/sufficient motivation for introducing categories

Part of the (human) reasoning happens at the level of categories

- Our goals can be *specific-instance-driven*

  E.g., I want to have dinner with my wife, I want to fuel my car, I want to

- Our goals can be *category-driven*

  E.g., I want to buy tomatoes, I'd like to read a (copy of) a book with title "AIMA"

Categories… do we need them?

We make predictions about the (future) evolution of the world using categories

- We recognize objects as belonging to certain categories
- We assign properties to the (objects belonging to) categories
- We assign models to the (objects belonging to) categories
- We then make "predictions"

E.g.: I am able to drive "any" car; I am able to use any fork and knife (at restaurant)

Categories and objects: how to represent them?

Objects: it is safe to assume there exists always a way to identify each of them, a unique name

E.g.: My car has a plate, I have a fiscal code, every pc on the internet has an IP address

Notice: this is not always true, there is part of the world that defies this assumption

In FOL, two possible alternatives:

1. categories as predicates
   e.g.: Car is a predicate that represents the category of cars;
   Car(aa123bb) means that the vehicle with plate aa123bb is a car

How to represent that (every) Car is a special type of (every) Vehicle?

Sub(Car, Vehicle) ... Not practical! it's a second order logic

Categories and objects: how to represent them?

In FOL, two possible alternatives:

1. categories as predicates
2. through **reification**: categories as objects as well
   E.g.: Member(aa123bb, Car) or $aa123bb \in Car$

What we would like to express about categories and objects? Is reification good enough?

- membership
- subclass: categories are organized in hierarchies, and subcategories inherit properties of the parents
  Note that the subclass relation generates taxonomies

Categories and objects, and reification

What we would like to express about categories and objects?

- **membership**

$$aa123bb \in Car$$

- **subclass**

$$Car \subset Vechicle$$

- members of a category enjoy some property (necessity)

$$(x \in \ Car) \Rightarrow hasWheels(x)$$

- members of a category can be recognized by some properties (sufficiency)

$$hasPlate(x) \wedge hasWheels(x) \Rightarrow x \in Car$$

- categories themselves can enjoy properties

  $Car \in VehicleType$ VehicleType is a category of categories…

Categories vs. Categories: Disjointness

Categories relate each other because of subclass relation…

… what else?

E.g.: Linneus stated that every animal could not be a vegetable, and the opposite. We would say that Animals and Vegetables are two disjointed categories, although they are subclass of LivingThing.

Formally:

$$Disjoint(s)$$
$$\Leftrightarrow$$
$$(\forall c_1, c_2 \quad c_1 \in s \ \wedge \ c_2 \in s \ \wedge \ c_1 \neq c_2 \ \Rightarrow \ c_1 \cap c_2 = \emptyset)$$

E.g.: Disjoint({Animals, Vegetables}).

Categories vs. Categories: Exhaustive Decomposition

Subcategories might "cover" all the possible instances of the parent category.

E.g.: Every one in this classroom is a student, or a professor (or both). "Student" and "Professor" categories provide an exhaustive decomposition of the category "People in this classroom"

Formally, given a category c, and a set s of categories, *s is an exhaustive decomposition of c if :*

$$ExhaustiveDecomposition(s, c)$$
$$\Leftrightarrow$$
$$(\forall i \quad i \in c \quad \Leftrightarrow \quad \exists c_2 \quad c_2 \in s \quad \land \quad i \in c_2)$$

E.g.: ExhaustiveDecomposition({Student, Professor}, PeopleInThisRoom).

Categories vs. Categories: Partition

If a category can be decomposed in more categories such that:

- they form an exhaustive decomposition
- they are disjoint

Then we have a partition

E.g.: Linneus stated that every living thing would partitioned into three subcategories (nowadays five categories).

Formally, given a category c, and a set s of categories, *s is a partition of c if* :

$$Partition(s, c) \\ \Leftrightarrow \\ Disjoint(s) \land ExahustiveDecomposition(s, c)$$

Physical Composition

Physical Composition

Objects can be made of other objects… the relation between an object and its parts is called **meronymy**: some objects (meronyms) are part of a whole (holonym).

E.g.:

- A car is composed of an engine, four wheels, five seats,…
- A laptop is made of a motherboard, a cpu, a keyboard, a screen, etc.
- My shopping bag contains three pieces of bread

The (philosophical) discipline that focuses on parts, on the relations among them, and relations with the whole, is named mereology.

Physical Composition

A rough, quite general distinction in the physical composition stems from the answer to this question: is there any **structural relation** between the parts, and between the parts and the whole? (other than the composition itself?)

- Answer **yes**: the relation is usually named **PartOf**
  E.g.: PartOf(engine123,car_aa123bc), PartOf(cylinder1345, engine123)

- Answer **no**: the relation is usually named **BunchOf**
  E.g.: BunchOf({Bread1, Bread2, Bread3})

Physical Composition – PartOf

The PartOf relation enjoys some properties:

- transitivity
$$PartOf(x,y) \land PartOf(y,z) \Rightarrow PartOf(x,z)$$
- reflexivity
$$PartOf(x,x)$$
always holds.

Physical Composition – BunchOf

- The BunchOf relation aims to define objects in terms of composition of other *countable* objects…
- … however, a structure in the composition relation is missing

E.g.: boxOfNails17 = BunchOf({nail1, nail2, nail3, nail7})

Why not use Set?

- Set is a mathematical concept
- Properties in general cannot be propagated to the Set concept

  E.g.: the weight of the box of nails is given by the sum of the nails contained in the box… the weight property of each nail has a relation with the weight of the box.

Physical Composition – BunchOf

- The BunchOf relation aims to define objects in terms of composition of other *countable* objects…
- … however, a structure in the composition relation is missing

E.g.: boxOfNails17 = BunchOf({nail1, nail2, nail3, nail7})

Why not use Set?

- Set is a mathematical concept
- Properties in general cannot be propagated to the Set concept

  E.g.: the weight of the box of nails is given by the sum of the nails contained in the box… the weight property of each nail has a relation with the weight of the box.

Measures

Measures

Among the many types of properties, objects can have measures, and specifically:

- **quantitative measures**: something that can be measured in a quantitative manner, w.r.t. some unit
  E.g.: the duration of this lesson, measured in seconds; the dimension in the space of a table (W x L x H);
- usually represented using a *unit function* that takes in input a number
  E.g.: length(table1) = cm(80)
- properties expressed in quantitative measures can/usually propagate upwards/downwards in the PartOf/BunchOf relations
  E.g.: the total weight of a car is given by the sum of the weights of its parts

Measures

Among the many types of properties, objects can have measures, and specifically:

- **qualitative measures**: something that can be measured using terms, and these terms come with an total/partial order relation
  E.g.: this lessons is {devastatingly boring, just boring, puzzling, amusing, fantastic} on a boring scale:
  implicitly: [dev.boring < boring < puzzling < amusing < fantastic]
- they do not propagate downward/upward
  E.g.: two boring lessons in the same day do not make the day twice boring, or devastatingly boring, or boring^2…
- the presence of an order relation however allow their use in object/category definitions

Qualitative Measures and Fuzzy Logic

What if we really need to make quantitative reasoning over qualitative measures?

Fuzzy Logic provides a (many-values-truth) semantics to qualitative measures

- many semantics, Lukasiewicz's one being among the more mentioned
- semantics focuses on the meaning of standard logical connectives (AND, OR, NOT)…
- … and on the combination of rules

Why not a probabilistic approach? Probabilistic logic and fuzzy logic capture different aspects of human reasoning…

Things or Stuffs?

Things or Stuff?

Till here, categories and objects were about things that can be individuated, or distinct from each other

- we can spot the car of the professor
- we can spot a single apple in the bunch of apples
- we can spot the people in this room, each one is an individual

Whether explicitly or implicitly:

- there is some **individuation criteria**
- mainly referred to the PartOf/BunchOf relations
- mainly related to the substance of the object itself
- related with the logical minimization criteria

Things or Stuff?

There are objects in our real world that defy the individuation criteria. In English:

- count nouns: car, apple, human
- mass nouns: water, air, butter

In the Cyc ontology, and Russel&Norvig as well, they adopt the distinction between:

- Things
- Stuff

Beware: a glass of water is different from water…

Things or Stuff – which one to use, and when?

A solution comes when looking at the properties

- some properties are referred to the substance of the object
  E.g.: the water boils at 100°C
  Such properties are named **intrinsic**, and they are retained even when some division is applied
- some properties refer to the object, and to the object structure relation with its parts
  E.g.: the weight of professor's car change if we take out a wheel
  Such properties are named **extrinsic**

Things or Stuff – extrinsic and intrinsic properties

A glass of water:

- has a volume
- has a weight
- is made of two "parts": the glass itself, and the "water within", each one with a different weight and volume (???)
- if we divide a glass of water, we will not obtain two glass of waters

The water:

- we can (infinitely?) divided it, and still it will be water
- it boils at a certain temp
- we can split it, and it will boil at the same temp

Things or Stuff – which one to use, and when?

A solution comes when looking at the properties

- a category of objects that will include in its definition only intrinsic properties is a substance; **Stuff** is the most general substance category
- a category of objects that will include in its definition any extrinsic property is a count noun; **Thing** is the most general object category

(such approach is the one adopted in Cyc, and by the textbook AIMA we are referring to in this course)

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

federico.chesani@unibo.it