Laboratory for Applied Ontology — Institute of Cognitive Science and Technology, Italian National Research Council

Research Institute for Computer Science of Toulouse - CNRS

Formal Ontology for Semanticists

Stefano Borgo, Nicola Guarino, Laure Vieu

Thanks to the LOA people and to Achille Varzi

www.loa-cnr.it

Course Outline

1. Ontology and natural language; ontology and ontologies; mereology (Guarino)
2. Ontology and formal NL semantics (Vieu)
3. Formal ontology and the space of foundational choices (Borgo)
4. Time, and space (Vieu)
5. OntoClean, DOLCE, and their applications to linguistic projects (Guarino)

Outline of this lesson

- Ontology and semantics
- Ontological commitment of natural language
- Ontology and ontologies
- A formal notion of ontology quality
- The role of formal ontology
- A glimpse of mereology

Ontology and semantics

- Strictly intertwined: ontology is about ***what there is***, semantics is about ***referring*** to what there is...
- Structural semantics vs. ***referential semantics***
- Different aspects of language, different roles of ontology *(lesson 2)*
  - Complex sentences (conjunctions, conditionals...)
  - Primitive sentences (predication)
  - Quantifiers and modifiers
  - Prepositions
  - Nouns and verbs
  - Discourse structure

The ontological commitment of natural language

- Every natural language (or maybe every contextualized sentence) *commits* to some ontology, in two ways [***theme of lesson 2***]:
  - Through a *closed* system of grammatical features
  - Through an *open* system of lexemes
- "Ontological semantics" [Nirenburg & Raskin 2004]: the semantics is driven by an ontology.
  - Practical role of ontologies for NLP systems

Which ontology for NL?

- Quine: every (logical) theory commits to the class of entities it ***quantifies on***.
- Problems:
  - Should every common noun correspond to an ontological category?
    - Questionable entities: Events, features, qualities, fictional characters...
  - Should different linguistic behaviors mark/reflect different ontological categories?

Descriptive vs. Revisionary Approach

- *Descriptive:* semantic structure of sentences is preserved (as best as possible)
- *Revisionary:* ontological eliminativism based on ***paraphrasability***:
  - John gives a kiss to Mary (Mary is given a kiss by John)
  - John kisses Mary (Mary is kissed by John)
  - John gives a flower to Mary
  - *John flowers Mary
  - There is a hole in this wall
  - This wall is holed
  - This statue has a long nose
  - This statue is long-nosed

The traps of revisionism

- Is ***systematic*** paraphrasing really possible (also for complex sentences)?
  - There are 7 holes in this piece of cheese
- How to choose ***whether*** paraphrasing?
  - Mary makes a leap
  - Mary makes a cake
- Can we account for ***proper inferences***?
  - There are two things John gave to Mary: a kiss and a flower
- ***Where to stop*** while eliminating entities?
  - Should we paraphrase everything in terms of bunches of molecules moving around? [not very interesting for a linguist...]

The rich ontology of Natural Language

Multiple *co-located events*
- ***John sings while taking a shower***

Multiple *co-located objects*
- I am talking here
- *This bunch of molecules is talking
- *What's here now is talking
- This statue is looking at me
- *This piece of marble is looking at me
- This statue has a strange nose
- *This piece of marble has a strange nose

Individual ***qualities***
- The temperature of this room is increasing
- I like the color of this rose
- The color of this rose turned from red to brown in one week

Ontology and polysemy

- Systematic polysemy [Pustejovsky]
  - Book: text/physical object
  - Window: opening/artifact
  - Apple: fruit/substance
  - ....
- A reason for not taking ontological semantics seriously? [Wliks]
- A reason for making clear the separation between lexicon and ontology? [Niremburg]
- A linguistic phenomenon *explained* by ontological dependence?

The role of *formal* ontology

Once the basic commitments are acknowledged, how to characterize (slight) differences in ontological assumptions?

From Ontology to Ontologies

- Ontology: the philosophical discipline
  - Study of what there (possibly) is
  - Study of the nature and structure of reality
    - Domain of entities
    - Categories and relations
    - Characterizing properties
- *An* ontology: a theoretical or computational artifact
  - "An explicit and formal specification of a conceptualization" (Gruber)
  - A specific artifact expressing the intended meaning of a vocabulary in terms of the nature and structure of the entities it refers to

What is an ontology?

- An artifact designed to account for the commitment of a language to a certain conceptualization
- A shared base vocabulary plus a formal characterization of its intended meaning
  - Vocabulary: concepts (categories) and relations
  - Characterization of meaning: *axioms* (logical formulas)
    - inferential behavior
    - restriction of models
- What counts is the structure of necessarily true knowledge
  - Describes no contingent situation (no assertional data)
  - The chosen vocabulary is arbitrary (P_1, P_2… would do)

What is a conceptualization

- Formal structure of (a piece of) reality *as perceived and organized by an agent, independently* of:
  - the ***vocabulary*** used
  - the actual occurence of a specific ***situation***
- Different situations involving same objects, described by different vocabularies, may share the same conceptualization.

Relations vs. Conceptual Relations

ordinary relations are defined on a ***domain*** D:

$$r_n \in 2^{D^n}$$

conceptual relations are defined on a ***domain space*** <D, W>

$$\rho_n : W \rightarrow 2^{D^n}$$ *(Montague's intensional logic)*

A ***conceptualization*** for D is a tuple **C** = <D, W, ℜ>, where ℜ is a set of conceptual relations on <D, W>

A ***model*** for a language **L** with vocabulary V is a structure

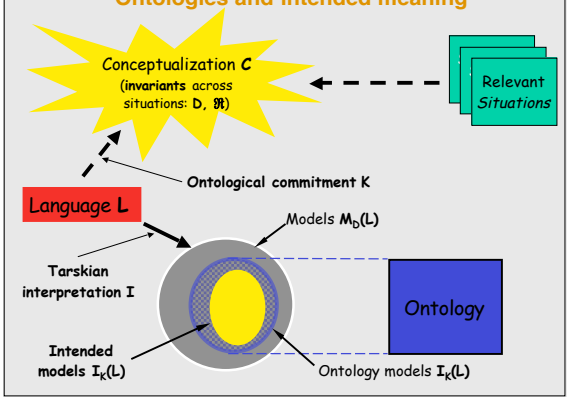
<**S**, I>, where **S** = <D, **R**> is a ***world structure*** and I: V ->D∪**R** is the usual interpretation function.

A model fixes a particular extensional interpretation of the language. Analogously, we can fix an ***intensional*** interpretation by means of a structure

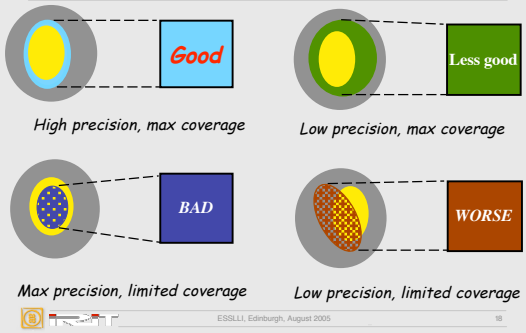
<**C**, ℑ>, where **C** = <D, W, ℜ> is a conceptualization and ℑ: V ->D∪ℜ is an *intensional interpretation function*.

We call such a structure **K**=<**C**, ℑ> an ***ontological commitment*** for **L**.
L ***commits*** to **C** by means of **K**.
C is the ***underlying conceptualization*** of **K**.

Ontologies and intended meaning



Ontology Quality: Precision and Coverage

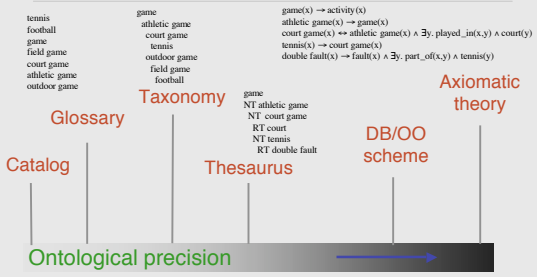


A quantitative metric for ontology coverage and precision

- Assumption: finite **D**, finite **W** *(examples)*

- Coverage = card($I_k \cap O_k$)/card(I_k)
- Precision = card($I_k \cap O_k$)/card(O_k)

Levels of Ontological Precision



When precision is not enough

Only one binary predicate in the language: on

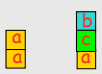
Only blocks in the domain: a, b, c, …

Axioms (for all x,y,z):

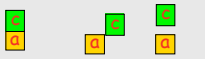
on(x,y) -> ¬on(y,x)

on(x,y) -> ¬∃z (on(x,z) ∧ on(z,y))

Non-intended *models* are excluded, but the rules for the competent usage of on in different *situations* are not captured.



Excluded situations



Indistinguishable situations

Precision vs. Accuracy

- In general, a single intended *model* may not discriminate among relevant alternative *situations because of*
  - Lack of *primitives*
  - Lack of *entities*
- Capturing all intended models is not sufficient for a "perfect" ontology

  Precision: non-intended *models* are excluded

  Accuracy: non-intended *situations* are excluded

Measuring ontological accuracy
(wrt benchmark examples)

- *Anomalous intended models* (set A_k): those that collapse intended and non-intended situations

$$\text{Accuracy} = (\text{card}(I_k)-\text{card}(A_k))/\text{card}(I_k)$$

Formal Ontology

- Theory of ***formal distinctions and connections*** within:
  - entities of the world, as we perceive it (***particulars***)
  - categories we use to talk about such entities (***universals***)
- Why ***formal***?
  - Two meanings: *rigorous* and *general*
  - Formal logic: connections between truths - neutral wrt ***truth***
  - Formal ontology: connections between things - neutral wrt ***reality***

Ontologies vs. Knowledge Bases

- Knowledge base
  - Assertional component
    - **reflects** ***specific (epistemic) states of affairs***
    - designed for ***problem-solving***
  - Terminological component (*ontology*)
    - ***independent*** of particular ***states of affairs***
    - Designed to support ***terminological services***

Ontological formulas are (assumed to be) ***necessarily true***

Formal Ontological Analysis

- **Theory of Essence and Identity**
- **Theory of Parts (Mereology)**
- **Theory of Wholes**
- **Theory of Dependence**
- **Theory of Composition and Constitution**
- **Theory of Properties and Qualities**

The basis for a common ontology vocabulary

Where everything starts: Mereology

- A bit of history
  - Lesniewski 1927-1931, *On the Foundations of Mathematics*
  - Greek *meros*
  - Alternative to Set Theory for escaping Russell's paradox
    - "the class of classes that are not members of themselves"
    - Expressed in a special logical language of its own "Ontology"
- Link with algebra: Tarski 1935
- The calculus of individuals, "nominalism"
  - Leonard & Goodman 1940
  - Expressed in first-order logic
  - No null individual
  - No abstract entities, no hierarchical distinction between individuals: a single relation of parthood
- Contemporary studies: Peter Simons (1986), Achille Varzi (1996)
  - All ontologies use a parthood relation, in the best cases fully specified with respect to Simons's work

Hasse diagrams

- Graph on finite domains
- Convention: all arcs vertical or oblique, implicitly oriented from bottom to top, strict order

x<y

Basic Mereology: M

- P, partial order (M_0):
  - (M1) $\forall x\, P(x,x)$
  - (M2) $\forall xyz\, ((P(x,y) \wedge P(y,z)) \rightarrow P(x,z))$
  - (M3) $\forall xy\, ((P(x,y) \wedge P(y,x)) \rightarrow x=y)$
- Weak supplementation (M):
  - (M4) $\forall xy\, (PP(x,y) \rightarrow \exists z\, (PP(z,y) \wedge \neg z=x))$

- Definitions
  - $PP(x,y) =_{df} P(x,y) \wedge \neg P(y,x)$
  - $O(x,y) =_{df} \exists z\, (P(z,x) \wedge P(z,y))$
  - $PO(x,y) =_{df} O(x,y) \wedge \neg P(x,y) \wedge \neg P(y,x))$

Extensional Mereology

- Extensionality
  - (E1) $\forall xy\, ((\exists z\, PP(z,x) \wedge \forall z\, (PP(z,x) \leftrightarrow PP(z,y))) \rightarrow x=y)$
  - (E2) $\forall xy\, (\forall z\, (O(z,x) \leftrightarrow O(z,y)) \rightarrow x=y)$
- Strong supplementation
  - (M5) $\forall xy\, (\neg P(y,x) \rightarrow \exists z\, (P(z,y) \wedge \neg O(z,x)))$

- Theorems
  - M_0+(M5) $\vdash$ (M4)
  - M_0+(M5) $\vdash$ (E1); M_0+(M5) $\vdash$ (E2)

Closure Mereology - 1

- Product
  (M6) $\forall xy\ (O(x,y) \rightarrow \exists z\ \forall t\ (P(t,z) \leftrightarrow (P(t,x) \wedge P(t,y))))$
  - z is *the* product of x and y, noted x•y

- Sum
  (M7) $\forall xy\ \exists z\ \forall t\ (O(t,z) \leftrightarrow (O(t,x) \vee O(t,y)))$
  - (E2) entails the unicity of z
  - z is the sum of x and y, noted x+y

Closure Mereology - 2

- Difference
  (M8) $\forall xy\ (\exists z\ (P(z,x) \wedge \neg O(z,y)) \rightarrow \exists z\ \forall t\ (P(t,z) \leftrightarrow (P(t,x) \wedge \neg O(t,y))))$
  - z is the difference of x and y, noted x-y
- Complement
  - Existence of the universe, noted U

  (M9) $\exists x\ \forall y\ P(y,x)$
  - Definition of the complement: $\sim x = U-x$ (exists for all $x \neq U$)

Classical/General Extensional Mereology

- General fusion
  (M6') $\exists x\ \phi(x) \rightarrow \exists z\ \forall y\ (O(y,z) \leftrightarrow \exists x\ (\phi(x) \wedge O(y,x)))$
  - Axiom schema, useful for infinite domains
  - unicity guaranteed by (E2), z is noted $\sigma x\ \phi(x)$
- Russell's description operator ι often used
  - $\sigma x\ \phi(x) = \iota z\ \forall y\ (O(y,z) \leftrightarrow \exists x\ (\phi(x) \wedge O(y,x)))$
- Sum, product and complement as fusions
  - $x+y = \sigma z\ (P(z,x) \vee P(z,y))$
  - $x•y = \sigma z\ (P(z,x) \wedge P(z,y))$
  - $\sim x = \sigma z\ (\neg O(z,x))$
- Universe
  - $U= \sigma x\ P(x,x)$
- NB: no null element!

Questioning Classical Extensional Mereology

- Problems with extensionality
  - Loosing or acquiring parts: identity across time
  - Identity between my body and the collection of my organs
- Problems with sums:
  - my nose and Caesar's toe;
  - Fusion: even stranger scattered infinite sums
  - First move: mereotopology to identify "wholes"
- Problems with transitivity
  - My hand is part of me, I'm part of my lab, but my hand is not part of my lab
  - The handle is part of the door, the door is part of the house. Is the handle part of the house?

Distinguishing various Part-Whole relations

- Linguistic and psychological evidence
  - Lyons 1977, Cruse 1986, Winston et al. 1987…
- Part-whole relations and *meronomies*
- A set of relations
  - Member / collection
    - This cow / the herd, John / the orchestra
  - Sub-collection / collection
    - Benelux / EU (but not USA / NATO)
  - Component-Integral Whole
    - The handle / the door, the engine / my car
  - Portion-Whole
    - A piece of cake
  - Substance-Whole
    - Some sugar / this cake
  - Piece-Whole
    - The left half of this table