Natural Language, Ontologies, Coherence

Christophe Fouqueré

(with M. Abrusci and M. Romano)
Laboratoire d'Informatique de Paris-Nord
Université Paris 13 - CNRS 7030

What does a linguistic concept denote?
What does a syntactic or semantic type denote?

What kind of coherence is there between elements of such denotations?

What are the consequences of using types/formulas?

Concepts and Types

➥Types as kinds of tags used in linguistic formal theories:

- Noun, Phrase, Verb, ...
- *e* and *t* (for individuals and truth-values)

➥Types used to analyze, to control inferences.

➥Two terms with same type should be in some sense interchangeable: their 'duals' are mutually *acceptable* contexts.

And duals of such a set of contexts should define a type.

Concepts and Types

➥ Concepts in *Linguistics*:

- ...
- Grammar: tense, aspect, mood, modality, ...
- Syntax: phrase, clause, grammatical function, grammatical voice
- Semantics, Pragmatics

➥ Concepts in *Natural Language*: *Being*, ..., *Table*, ...

➥ "A conceptualization is an abstract view of the world."

Types and Linear Logic

Categorial Grammar is widely used, as such or in variants, as it relates Natural Language as a typed functional language, hence to λ-calculus:

- linguistic information is encoded in the lexicon via the assignment of syntactic types to lexical items,
- expressions are either functions or arguments.

Curry-Howard correspondence allows to view linguistic theories as formulas in a suitable Logics.

Linear Logic extends the intuitionnistic approach:

- Full Linear Logic may be viewed as a strongly typed programming language
- Non-intuitionnism may be interpreted for example as exception handling
- Formulas may be interpreted as usable resources

Questions remain:

- What does a type denote?
- Is there any relation between elements of (denotations of) concepts and types?

The *Geometry of Interaction* program initiated by JY Girard tries to fully integrate syntax and semantics:
➥logical objects give the denotation of their use.

So let us look at ontologies and concepts ...

Ontologies

Quoting Guarino ("Handbook on Ontologies"):

“A body of formally represented knowledge is based on a conceptualization: the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them.

A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose.

Every knowledge base, knowledge-based system, or knowledge-level agent is committed to some conceptualization, explicitly or implicitly.”

Ontologies

Quoting Quine's slogan ("On what there is"):

"To *be* is to be the value of a bounded variable"

➥The logic to be adopted, according to Quine, is First Order Logic relying on set theory.

Hence:

- concepts and relations are denoted by sets of objects,
- data that are recorded in the system as instantiating those concepts and relations.

Ontologies

However:

- such a choice implies that any change in the extensional picture produces also a change of conceptualization
- It means that even the turn-over, over the time, of the instances of a concept causes an unending change of the reference conceptualization.

Ontologies

(works done with Abrusci and Romano)

The focus is on the extensional level, i.e. on “real” objects:

- relations among resources are encoded in a logical framework,
- hence the logical interpretation should rely on structures richer than sets: *Coherence Spaces*

The interpretation of a concept produces:

- graph theoretical objects
- the determination of the extensional counterpart within the collection of *resources*.

Ontologies

What is a resource?

- In concrete / web ontologies: data stored in some base, tags put by a user
- In Natural Language: words, sentences produced or heard, dialogues, ...

Ontologies

Formal Ontology:

- Set of concepts together with relations (or roles) between them
- Presented as a first-order theory:
  - A **concept** is a unary predicate
  - A **relation** (or **role**) is a binary predicate
  - An entity (or **individual**, or datum) is a constant
- Presented as two parts:
  - A **T-box**, the *terminology*: a set of axioms on concepts and relations (without constants)
  - A **A-box**, the *assertions*: a set of atomic axioms with constants (without variables)

Ontologies

- standard description languages (OWL and variants)
- reasoning by means of description logics or modal logics

Ontologies

- typical kinds of **inferences** on T-boxes: (all reducible to $C \sqsubset D$)
  - satisfiability: of a concept wrt a theory
  - subsumption: extension of C is included in extension of D
  - equivalence: equality of extensions
- typical kinds of **reasoning** with A- and T-boxes:
  - consistency: of the A-box wrt the T-box
  - retrieval: of instances of some concept
  - property expansion: given properties of an entity, infer other concepts this entity is instance of
- **querying** an ontology: queries by means of graph patterns (SPARQL)

Folsksonomies

Current developments: Web2.0

- give back to the Web its original nature, that of a networked platform where every node of the net is as important as any other
- differences with Web1.0:
  - Content Management Systems (structured by owner) vs Wikis (structured by community)
  - Directories (hence taxonomies) vs Tagging (hence folksonomies)

Folsksonomies: tagging

What is **(free) tagging**?

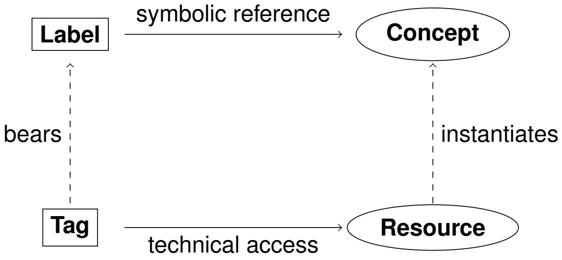
- Everybody can stick tags on everything!
- To put triples *User – Tag – Resource*

Folksonomies

In a first approximation, one may relate free tagging to ontology:

A **concept** defines a class of **objects** by means of some linguistical formulation, typically the **term** that signals the concept.

More precisely (Monnin, Romano):



Folksonomies

Two (broad) types of use:

- Users tag references to resources discovered on the web, and may share their taggings
- Users tag resources (their own, and may also add tags to other users' resources)

Folksonomies

Note that:

- "Semantics" of tag is user-dependent
- Tag names may be useless for understanding their usage

What is important is the whole tagging mechanism, not necessarily at first names used for tags or labels.
Even if semantics of tag names or labels may be more and more precise, the further they are used.

➥Hence necessity to take into account *what* is related by tags.

Ontologies / Folksonomies

Needs for a change the viewpoint:

- *from* a predefined, typed, uniform, perspective,
- *to* a (maybe) post-definable, a priori untyped, subjective (in some way) perspective

- *from* immutable data, concepts and relations,
- *to* interpretations subject to variation

Coherence Spaces

Coherence Spaces are defined as a denotational semantics for Linear Logic.

Definition

A coherence space A is a countable graph with vertices $|A|$ and a coherence relation $\frown_A$ reflexive and symmetric.

- A propositional letter is denoted by a coherence space.
- Connectives are denoted by operations on coherence spaces.

What results?

- A proof is denoted by a clique.
- A (multiplicative) proof structure (formulas, axioms, cuts) is a proof iff its *experiments* are coherent wrt the par of the conclusions.

Coherence Spaces: operations

Definition

- $\mathbf{A}^{\perp}$ is defined such that $|A^{\perp}| = |A|$
  and $x \frown_{A^{\perp}} y$ iff $x = y$ or $x \not\frown_A y$
- $\mathbf{A} \otimes \mathbf{B}$ is defined such that $|A \otimes B| = |A| \times |B|$
  and $(x, x') \frown_{A \otimes B} (y, y')$ iff $x \frown_A y$ and $x' \frown_B y'$
- $\mathbf{A} \multimap \mathbf{B}$ is defined such that $|A \multimap B| = |A| \times |B|$
  and $(x, x') \frown_{A \multimap B} (y, y')$
  iff $(x \frown_A y$ then $x' \frown_B y'$ and $x \neq y$ then $x' \neq y')$
- $\mathbf{A} \oplus \mathbf{B}$ is defined such that $|A \oplus B| = \{0\} \times |A| \cup \{1\} \times |B|$
  and $(0, x) \frown_{A \oplus B} (0, x')$ iff $x \frown_A x'$,
  $(1, y) \frown_{A \oplus B} (1, y')$ iff $y \frown_B y'$,
  $(0, x) \not\frown_{A \oplus B} (1, y)$.

Coherence Spaces and Ontologies

An *ontology* is a triple (O, M, Φ) such that:

- O is a set of **predicate** and **relation** symbols
- M is a set of **individuals** (or constants)
- Φ is defined on O such that:
  - $\Phi(P) \subset M$: each predicate symbol is associated to a set of individuals,
  - $\phi(R) \subset (M \times M)$: each relation symbol is associated to a set of pairs of individuals.

In the following, we restrict ourself to *decidable* ontologies.

Coherence Spaces and Ontologies

We build the coherence space forgetting about what means *a priori* individuals, predicates and relations.

Definition

An *ontological compatibility space* (OCS) $\mathfrak{O}$ defined on a KB (O, M, Φ) is a coherence space such that:

- $|\mathfrak{O}| = \bigcup_{o \in O} \Phi(o)$
- $x \frown_{\mathfrak{O}} x$
- $x \frown_{\mathfrak{O}} \langle x', y' \rangle$ and $\langle x, y \rangle \frown_{\mathfrak{O}} x'$.
- $x \frown_{\mathfrak{O}} y$ when $\exists P \in O, \{x, y\} \subset \Phi(P)$
- $\langle x, y \rangle \frown_{\mathfrak{O}} \langle x', y' \rangle$ iff $\exists R \in O, \{\langle x, y \rangle, \langle x', y' \rangle\} \subset \Phi(R)$

Coherence Spaces and Ontologies
Example

Let us consider the following example in RDF (Romano):

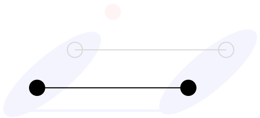
```
<rdf:RDF
   xmlns:rdf="http://www.w3.org/1999/02/22--rdf--syntax--ns#"
   xmlns:rdfs="http://www.w3.org/2000/01/rdf--schema#"
   xmlns:foaf="http://xmlns.com/foaf/0.1/">
   <foaf:Personrdf:ID="me">
      <foaf:name>MarcoRomano</foaf:name>
      <foaf:workInfoHomepagerdf:resource
                                 ="http://logica.uniroma3.it/~romano"/>
      <foaf:mboxrdf:resource="mailto:m.romano@uniroma3.it"/>
      <foaf:knows>
         <foaf:Person>
         <foaf:name>V.MicheleAbrusci</foaf:name>
         <foaf:mboxrdf:resource="mailto:abrusci@uniroma3.it"/>
         </foaf:Person>
      </foaf:knows>
   </foaf:Person>
</rdf:RDF>
```

Coherence Spaces and Ontologies

Example

The interpretation is the following one:

```
<rdf:RDF
  ...
  <foaf:Personrdf:ID="me">
    <foaf:name>M. Romano</foaf:name>
    <foaf:workInfoHomepagerdf:resource
        ="http://logica.uniroma3.it/~romano"/>
    <foaf:mboxrdf:resource
        ="mailto:m.romano@uniroma3.it"/>
    <foaf:knows>
      <foaf:Person>
        <foaf:name>M. Abrusci</foaf:name>
        <foaf:mboxrdf:resource
            ="mailto:abrusci@uniroma3.it"/>
      </foaf:Person>
    </foaf:knows>
  </foaf:Person>
</rdf:RDF>
```

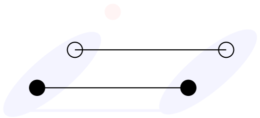


2 Persons,

Coherence Spaces and Ontologies
Example

The interpretation is the following one:

```
<rdf:RDF
  ...
  <foaf:Personrdf:ID="me">
    <foaf:name>M. Romano</foaf:name>
    <foaf:workInfoHomepagerdf:resource
        ="http://logica.uniroma3.it/~romano"/>
    <foaf:mboxrdf:resource
        ="mailto:m.romano@uniroma3.it"/>
    <foaf:knows>
      <foaf:Person>
        <foaf:name>M. Abrusci</foaf:name>
        <foaf:mboxrdf:resource
            ="mailto:abrusci@uniroma3.it"/>
      </foaf:Person>
    </foaf:knows>
  </foaf:Person>
</rdf:RDF>
```



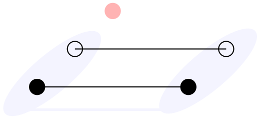
2 Persons, 2 Emails,

Coherence Spaces and Ontologies

Example

The interpretation is the following one:

```
<rdf:RDF
  ...
  <foaf:Personrdf:ID="me">
    <foaf:name>M. Romano</foaf:name>
    <foaf:workInfoHomepagerdf:resource
        ="http://logica.uniroma3.it/~romano"/>
    <foaf:mboxrdf:resource
        ="mailto:m.romano@uniroma3.it"/>
    <foaf:knows>
      <foaf:Person>
        <foaf:name>M. Abrusci</foaf:name>
        <foaf:mboxrdf:resource
            ="mailto:abrusci@uniroma3.it"/>
      </foaf:Person>
    </foaf:knows>
  </foaf:Person>
</rdf:RDF>
```

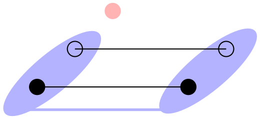


2 Persons, 2 Emails, 1 Web page,

Coherence Spaces and Ontologies
Example

The interpretation is the following one:

```
<rdf:RDF
  ...
  <foaf:Personrdf:ID="me">
    <foaf:name>M. Romano</foaf:name>
    <foaf:workInfoHomepagerdf:resource
        ="http://logica.uniroma3.it/~romano"/>
    <foaf:mboxrdf:resource
        ="mailto:m.romano@uniroma3.it"/>
    <foaf:knows>
      <foaf:Person>
        <foaf:name>M. Abrusci</foaf:name>
        <foaf:mboxrdf:resource
            ="mailto:abrusci@uniroma3.it"/>
      </foaf:Person>
    </foaf:knows>
  </foaf:Person>
</rdf:RDF>
```



2 Persons, 2 Emails, 1 Web page, 2 pairs Person-Email

Coherence Spaces and Ontologies

We remark the following immediate points:

- Any concept of an ontology, i.e. the extension of a predicate or a relation, is a **clique** of the corresponding OCS.
- The empty set is a clique: it denotes an "impossible" concept.
- Each clique may be interpreted as a *potential* concept.

Coherence Spaces and Ontologies

One may ask which cliques may *really* be concepts: maximal cliques ?

A *Maximal clique C* may be defined in different ways:

- clique of elements such that two individuals share a *common* property:
  IF x is s.t. $\forall y \in C, y \neq x, \exists P$ s.t. $\{x, y\} \subset \phi(P)$ THEN $x \in C$
- clique of elements such that a *set* of properties is shared by all individuals:
  IF x is s.t. $\forall P \in \mathcal{P}, x \in \phi(P)$ THEN $x \in C$
- ...

However, it may be more useful to consider that 'concepts' are defined by interaction with user requests.

Coherence Spaces and Ontologies

Operations

Standard operations and relations on ontologies (1):

- An ontology $\mathcal{O}$ is a **segment** of an ontology $\mathcal{O}'$ if $\mathcal{O}$ is a restriction of $\mathcal{O}'$ to a part of its language
- An ontology $\mathcal{O}$ **inherits** from an ontology $\mathcal{O}'$ wrt a language L if the theory of $\mathcal{O}'$ restricted to L is included in the theory of $\mathcal{O}'$

Coherence Spaces and Ontologies

Operations

Standard operations and relations on ontologies (2):

- **Union** of two (distinct) ontologies
- **Mapping** α from an ontology (O, M, ϕ) to another one (O', M', ϕ'):
  - If $a \in \phi(P)$, $\alpha(a)$, $\alpha(P)$ defined, then $\alpha(a) \in \phi'(\alpha(P))$
  - If $< a, b > \in \phi(R)$, $\alpha(a)$, $\alpha(b)$, $\alpha(R)$ defined, then $< \alpha(a), \alpha(b) > \in \phi'(\alpha(R))$
- **Refinement**, **Alignment** as special cases of mapping
- **Merging** of two ontologies (O, M, ϕ) and (O', M', ϕ') as a partial alignment and inheritance of what is not aligned.

Coherence Spaces and Ontologies

Operations

Operations on ontologies may be interpreted in terms of operations on OCS:

- The $\oplus$ operation on OCS corresponds to the union of two ontologies.
- The $\multimap$ operation on OCS corresponds to mapping ontologies:
  - A mapping α is represented as a clique in $\mathfrak{O} \multimap \mathfrak{O}'$
  - $P \times \alpha(P)$ and $R \times \alpha(R)$ are cliques of $\mathfrak{O} \otimes \mathfrak{O}'$

Coherence Spaces and Ontologies

Such a framework allows to relate also folksonomies to ontologies.

- 2 resources are in relation if they have some *quality* in common (maybe subjective)
- a tag, a concept is represented as a clique
- Note that the viewpoint may be changed: 2 tags are in relation if there exists a common resource, . . .
- What is a point? What is a coherence structure?
  - Mainly logical structures, i.e. proofs,
  - that may be questioned, i.e. reduced by cuts.
  - Hence **Ludics** or Game Semantics

Thanks for your attention