

Ontologies and Language — comment on Mark Musen's 'Differing views of biomedical ontology'

Pierre Zweigenbaum

STIM: Mission de recherche en sciences et technologies
de l'information médicale,
DSI, Assistance Publique – Hôpitaux de Paris ;
CRIM, INaLCO ;
U729, INSERM

pz@biomath.jussieu.fr
<http://www.biomath.jussieu.fr/~pz/>

1 Ontologies

- Different senses of the word “ontology”
- Different purposes
- Different levels of formality
- Different stages

2 Language and the world

- Shared meanings
- Language builds conceptualizations
- Language builds conceptualizations
- Natural vs formal language
- Language (texts) as an entry point to world knowledge

Background

- [Maths, Physics]
- Computational linguist
- A word on unicorns

Background

- [Maths, Physics]
- *Computational* linguist
computation = formal
- A word on unicorns

Background

- [Maths, Physics]
- Computational *linguist*
computation = formal language = informal
- A word on unicorns

Background

- [Maths, Physics]
- Computational linguist
computation = formal vs language = informal
⇒ no instance in the world?
- A word on unicorns

My neighbor the unicorn (Cluny Museum)



1 Ontologies

- Different senses of the word “ontology”
- Different purposes
- Different levels of formality
- Different stages

2 Language and the world

- Shared meanings
- Language builds conceptualizations
- Language builds conceptualizations
- Natural vs formal language
- Language (texts) as an entry point to world knowledge

Different senses of the word “ontology”

I say ‘ontology’, you say ‘ontology’, but do we mean the same thing?

- Alexa: state your bias
- Robert: better define ‘ontology’, ‘terminology’

► Language

A serious issue

- (nearly) Nobody agrees to have ever seen any ‘proper ontology’
except self-referencing, which is to be discarded

A serious issue

- (nearly) Nobody agrees to have ever seen any ‘proper ontology’
except self-referencing, which is to be discarded
- No instance \Rightarrow is ‘ontology’ a universal?

A serious issue

- (nearly) Nobody agrees to have ever seen any ‘proper ontology’
except self-referencing, which is to be discarded
- No instance \Rightarrow is ‘ontology’ a universal?

Give me 1\$ each time you use the ‘O-word’

A serious issue

- (nearly) Nobody agrees to have ever seen any 'proper ontology'
except self-referencing, which is to be discarded

- No instance \Rightarrow is 'ontology' a universal?

Give me 1\$ each time you use the 'O-word'

- A new *dilemmaton*:
'believe in unicorns' OR 'ontologies do not exist'

A serious issue

- (nearly) Nobody agrees to have ever seen any ‘proper ontology’
except self-referencing, which is to be discarded

- No instance \Rightarrow is ‘ontology’ a universal?

Give me 1\$ each time you use the ‘O-word’

- A new *dilemmaton*:
‘believe in unicorns’ OR ‘ontologies do not exist’

► Language

Ontologies for different purposes

- Different purposes → different views on the world, need for different distinctions
- [Mark:]
 - to provide a classification of biomedical entities
 - to summarize and annotate data
 - to simplify the engineering of complex software systems
 - to mediate among different social groups
 - to mediate among different software components
 - to provide a formal specification of biomedical knowledge

Different purposes

- Some correspond to different socio-professional needs:
 - to provide a classification of biomedical entities
 - to summarize and annotate data
 - to simplify the engineering of complex software systems
- cf sociolects

Different purposes

- Some correspond to different socio-professional needs:
 - to provide a classification of biomedical entities
 - to summarize and annotate data
 - to simplify the engineering of complex software systems→ cf sociolects
- Some aim at bridging different socio-professional needs:
 - to mediate among different social groups
 - to mediate among different software components→ to which extent is this possible?

► Language

Different purposes

- Some correspond to different socio-professional needs:
 - to provide a classification of biomedical entities
 - to summarize and annotate data
 - to simplify the engineering of complex software systems→ cf sociolects
- Some aim at bridging different socio-professional needs:
 - to mediate among different social groups
 - to mediate among different software components→ to which extent is this possible?
- More open-ended purpose:
 - to provide a formal specification of biomedical knowledge→ reduce to previous purposes?

► Language

More or less formal ‘ontologies’

- Thesaurus, classification, nomenclature
- Use of (description) logics (consistency, inferencing capabilities): ‘formal ontology’, ‘upper-level ontology’

More or less formal ‘ontologies’

- Thesaurus, classification, nomenclature
- Use of (description) logics (consistency, inferencing capabilities): ‘formal ontology’, ‘upper-level ontology’
- To get formal:
 - Delimit metalanguage, with its primitives (*is-a*, *part-of*, ...)
 - Ensure they will only get one interpretation: **fix context**
 - → Specify purpose
- Can a formal ontology use words from a natural language?

Ontology sharing and reuse

- Reuse an existing 'ontology' for some other purpose
 - No examples of actual reuse
(human remodeling work always involved)
 - Can this be done automatically?
 - This needs an additional, external metalanguage
- Mediate between different ontologies
 - = reuse?
- Provide a universal, formal ontology to support purpose-oriented ontologies
 - Necessarily limited to purpose-independent things
 - Relations: *is-a*; *part-of*?
 - Entities: Top-level ontology? To which extent does it bear consequences on the organization of domain entities?

Different stages of one ontology

A Knowledge Engineering perspective

- Authoring a formal ontology: different stages towards a target ontology (TIA WG; Bachimont)
 - Facts: Description of domain through text corpus
 - Acquire terminology (terms), structure (relations) ▶ Language
 - Regional ontology (with natural language definitions)
 - Formal ontology (formalization)
 - Computational ontology (operationalization)
- Maintaining a formal ontology: interface with humans
 - GALEN: intermediate representation
 - Text-based authoring: link back to text (context)
 - Rely on natural language definitions for primitive concepts:
Use natural language as the metalanguage

- 1 Ontologies
 - Different senses of the word “ontology”
 - Different purposes
 - Different levels of formality
 - Different stages
- 2 Language and the world
 - Shared meanings
 - Language builds conceptualizations
 - Language builds conceptualizations
 - Natural vs formal language
 - Language (texts) as an entry point to world knowledge

Shared meanings

Meaning negotiation

- People understand each other because they can agree on shared meanings
- In case of lack of initial understanding they can 'negotiate' meaning

We must agree on shared meaning(s) for 'ontology'

Meaning drift

- philosophy
- computer science: formal ontology
- medical informatics: terminology, classification

Choose the right name(s)

Is this a problem with the choice of (a) good term(s)?

- *formal ontology* (Thomas: *top-level ontology*)
- *casual ontology* (Olivier)
- *informal ontology* (*material ontology* (Bachimont; Barry))
- *regional ontology* (Bachimont)
- *terminology, thesaurus*
- *classification, nomenclature*
- *terminological knowledge base* (Meyer; TIA WG)
- *termino-ontological product* (TIA WG)

Choose the right name(s)

Is this a problem with the choice of (a) good term(s)?

- *formal ontology* (Thomas: *top-level ontology*)
- *casual ontology* (Olivier)
- *informal ontology* (*material ontology* (Bachimont; Barry))
- *regional ontology* (Bachimont)
- *terminology, thesaurus*
- *classification, nomenclature*
- *terminological knowledge base* (Meyer; TIA WG)
- *termino-ontological product* (TIA WG)



Call for Académie Française?

Choose the right thing(s)

Is this a problem with the choice of the right thing(s) we want to talk about?

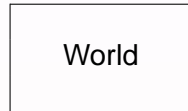
[Alan: people do not argue about the concepts, they argue about the names]

- Both need to be taken care of
- The two faces of the same coin?

Our grasp on the world

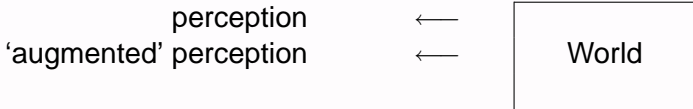
Look at the world through the eyes of language

perception



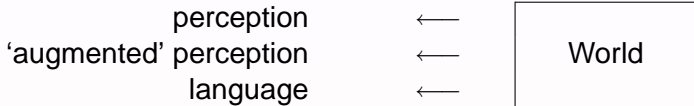
Our grasp on the world

Look at the world through the eyes of language



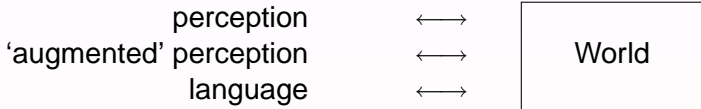
Our grasp on the world

Look at the world through the eyes of language



Our grasp on the world

Look at the world through the eyes of language



The arrows are bidirectional

Language builds conceptualizations

- Language shapes our view of the world

Language builds conceptualizations

- Language shapes our view of the world
- Not the same in all languages:
 - *laïcité* ('separation of state and church'?)
 - *blue/green* borderline in Chinese
 - *finding?*

Language builds conceptualizations

- Language shapes our view of the world
- Not the same in all languages:
 - *laïcité* ('separation of state and church'?)
 - *blue/green* borderline in Chinese
 - *finding?*
- 'Abstract' concepts: *freedom, love*

Language builds conceptualizations

- Language shapes our view of the world
- Not the same in all languages:
 - *laïcité* ('separation of state and church'?)
 - *blue/green* borderline in Chinese
 - *finding?*
- 'Abstract' concepts: *freedom, love*
- Social constructs, differ across social groups (*sociolects*)
 - Borrowing from one social group to another: *meaning drift*
 - philosophy: *ontology*
 - computer science (knowledge representation and knowledge engineering): *formal ontology*
 - medical informatics: *terminology, classification*
 - everybody: 'anything'

Natural vs formal language

- Natural language
 - expressive capacity
 - its own metalanguage (*this is metalanguage*)
 - importance of context
 - system of interdependent definitions
- Formal language
 - rigor
 - generative power
 - provability (?)
 - computability (?)
 - precisely defined constructs
- On which side is 'ontology'?

Language (texts) as an entry point to world knowledge

Study language use (text corpora) to know about the world

- Inventory of terms (entities)
- Inventory of relations
- Inventory of facts
- A text corpus must be built by specifying selection criteria:
 - domain
 - genre
 - ...

These criteria characterize the relevant social group and purpose (task)

(see presentations by Olivier and Christiane)