Experiences with developing DOLCE

Claudio Masolo

Laboratory for Applied Ontology, ISTC-CNR, Trento, Italy
masolo@loa.istc.cnr.it

Ontology Summit 2021, The Ontological Landscape – 10 March 2021

1

- DOLCE (2002)
  Descriptive **O**ntology for **L**inguistic and **C**ognitive **E**ngineering
  - Developed in the context of the WonderWeb, *Ontology Infrastructure for the Semantic Web* – EC project (2002-2004)
  - http://www.loa.istc.cnr.it/dolce/overview.html
- DOLCE-CORE (2009) - update of the core fragment of DOLCE
- Several extensions of DOLCE have been developed
- Ongoing ISO-standardization activity

2

- Basic taxonomy of DOLCE

Endurant (Object)
- Physical
  - Amount of matter
  - Physical object
  - Feature
- Non-Physical
  - Mental object
  - Social object

Perdurant (Event)
- Static
  - State
  - Process
- Dynamic
  - Achievement
  - Accomplishment

Quality
- Physical Quality
  - Spatial Location
  - ...
- Temporal Quality
  - Temporal Location
  - ...
- Abstract Quality

Abstract
- Quality Region (Region)
  - Time Region
  - Space Region
  - Color Region
  - ...

3

- Primitive relations of DOLCE
  - *Parthood simpliciter* - $\mathsf{P}(x, y)$ "x is part of y"
  - *Temporary parthood* - $\mathsf{tP}(x, y, t)$ "x is part of y at time t"
  - *Constitution* - $\mathsf{K}(x, y, t)$ "x constitutes y at time t"
  - *Participation* - $\mathsf{PC}(x, y, t)$ "object x participates in event y at t"
  - *Quality* - $\mathsf{qt}(x, y)$ "individual quality x inheres in y"
  - *Quale* - $\mathsf{ql}(x, y, t)$ "region x is the quale of individual quality y at t"
- ⋆ 74 axioms and 103 definitions

4

- WonderWeb Project
  - development of the *infrastructure* required for the large-scale deployment of ontologies as the foundation for the Semantic Web
  - establishment of Web standard ontology languages and the development of ontological engineering technology
  - ⋆ Languages + Ontologies + Ontology engineering + Tools and services

- *Toolkit* supplemented with a set of ontologies, covering a wide range of application domains that can be used to construct more detailed domain ontologies and to integrate existing ontologies

  ...which kind of ontologies?

5

- **O**ntology *vs*. **o**ntology
  Theory of *reality vs*. an arbitrary model

  Philosophers focus more on the nature of reality than on different *models* of reality (*being qua being vs*. *content qua content*)

  Computer scientists focus more on the nature of reasoning than on modeling (modeling is leaved to the end user)

- *Ontological promiscuity of AI*: any agent creates its own ontology based on its usefulness for the task at hand [Genesereth and Nilsson, 1987]

6

- ⓞntology - between **O**ntology and **o**ntology
  - *Pluralistic view* to avoid both a
    - – monolithic approach and
    - – arbitrary models

- It requires two shifts

7

- From **O**ntology to ⓞntologies
  - *Ontological*, *conceptual*, *cognitive*, *hypothetical*, etc. entities may all be relevant in knowledge representation and conceptual modeling
  - No absolute *reality*: different 'realities'/views are useful to account for natural language, common sense, etc.

∼ From **L**ogic to **l**ogics

  - No absolute *truth*: different truths underly different reasoning mechanisms that make sense in specific *contexts*

- Grounding **o**ntologies into ⓞntologies
  - ⓞntologies provide *conceptual handles* to carry out a coherent and structured analysis of the domains of interest that help to avoid *ad-hoc solutions*
  - ⓞntologies help to avoid arbitrary models and to build well founded models

9

- An ⓞntology is first of all for understanding each other but not necessarily for thinking in the same way: help recognizing and understanding *disagreements* as well as *agreements*
- An ⓞntology is philosophically/conceptually well founded and it explicitly characterizes its basic assumptions (*rich axiomatization*)
    - to avoid mis-using and misunderstandings
    - to be semantically transparent w.r.t the ontological commitment
- An ⓞntology has a large scope and it is reusable in different modeling scenarios

10

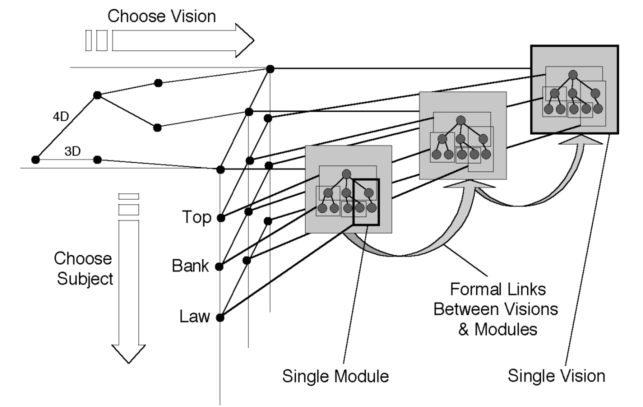
- The philosophical analysis is an important input
  - Focus on very basic relations (e.g. *identity*, *parthood*, *dependence*, *constitution*, *participation*), that are not specific to particular domains but can be used to characterize ontological distinctions and can be suitably refined to match specific requirements
- But *linguistic*, *cognitive science*, *mathematics*, etc. have also investigated general notions interesting from a modeling perspective (multidisciplinar approach)
  - Ex.: *qualities* and *qualia* in DOLCE are based on the theory of Conceptual Spaces of P. Gärdenfors and on linguistic studies

- How the plurality of ⓞntologies can be managed?

  Library - A (small) set of ⓞntologies carefully linked, justified, documented, and positioned with respect to a space of possible choices

  - Starting point for building new ontologies
    $\Rightarrow$ *re-using & modularization*
  - Reference point for easy and rigorous comparison among different ontological choices
    $\Rightarrow$ *semantic integration / partial inter-operability*
  - Furnishes a common framework for analyzing, harmonizing and integrating existing ontologies and metadata standards
    $\Rightarrow$ *trust*

12


Choose Vision
4D
3D
Choose
Subject
Top
Bank
Law
Formal Links
Between Visions
& Modules
Single Module
Single Vision

13

- WonderWeb library
  - DOLCE
  - BFO (axioms developed by P. Grenon in collab. with B. Smith and LOA)
  - OCHRE (developed by Luc Schneider)
- Modular approach
  - A modularization of DOLCE exists (done in CASL/HETS) and it has been used to prove its consistency
  - ⋆ COLORE repository (based on Common Logic) and DOL (Distributed Ontology, Modeling and Specification Language - OMG specification) are examples of frameworks (languages+tools+methodologies) supporting modularization

- The OntoCommons H2020 CSA project
(*Ontology-driven Data Documentation for Industry Commons*)
  - Standardisation of data documentation across all domains related to materials and manufacturing
  - Develop a ready-to-use Ontology Commons EcoSystem (OCES) for data documentation, including a set of ontologies and tools
  - ⋆ The ontological top-part of OCES is basically an (integrated) library of top-level ontologies that contains DOLCE, BFO, EMMO (the inclusion of other ontologies is under analysis)

- ⓞntologies are used in classical/hybrid AI, Knowledge Representation, and Semantic Web

  but they are also used

  - to clarify the semantics of *Conceptual Modeling* languages (e.g. UML) or extend them by introducing ontologically well founded primitives and 'design-patterns' (see OntoUML and Ontology Design Patterns)
  - to improve the ontological foundation of the lexicons (and to linguistically found the ontologies) via alignments with *computational lexicons* (see OntoWordNet)

- A rich axiomatization is an essential feature of ⊚ntologies
- Slogans
  - *Ontological analysis: study of content qua content*

    What must be modeled needs to be studied, understood, and analyzed *as such*, independently of the way it will be represented and used
  - *No ontology without ontological analysis!*

    First ontological analysis, *then* knowledge representation

★ Attempt to *separate* the ontological analysis from its formal representation and its use in applications

- Ontological analysis can be carried out in a *precise* but not necessarily *formal* way (it often happens in analytic philosophy)

- However, in my experience, formalization is extremely useful to
  - take into account very subtle distinctions, compare them, and check what are the consequences of specific choices
  - make explicit, objective, and communicable the analysis

- In my experience, logic and axiomatization play an important role at both the *analytical* level and *implementation* level

- The choice of the formal language is problematic
  - in analytical terms, a highly expressive language is desirable (better characterization of the primitives)
  - in applicative terms, high expressive power compromises computat. efficiency (no automatic inference and classification in reasonable time)

- WonderWeb: separate the ‘analytical’ from the ‘implementation’ language
  - Start from (modal) FOL that is well known and allows a good characterization of primitives and then *approximate* the theory in OWL which is less expressive but has good computational behavior

  But...

  - automatic translators are very difficult
  - manual translations can be difficult and time consuming
  - approximations change the (meaning of) primitives and are often unable to capture most of the analysis represented in the FOL theory

20

- Ontological commitment and formal language

- Some (kinds of) entities can be introduced in the domain of quantification to overcome limitations of the language (a relevant aspect in applications)
- *Quine*'s maxim: "*to be is to be a value of a variable*" (something exists iff it is in the domain of quantification)

⋆ It is not always simple to understand the 'true' ontological commitment of models developed under applicative constraints

- *Example* 1. n-ary relations in DLs

  $R(a_1, \ldots, a_n)$ is represented by introducing a state of affairs s such that

  - $R^*(s)$
  - $ARG_1(a_1, s) \wedge \ldots \wedge ARG_n(a_n, s)$

- Are states of affairs part of the ‘reality’?

- *Example* 2. Institutions/organizations

  Imagine to “re-express truths about the global political consequences of a decline in the GNP [Gross National Product] of Eastern Europe in terms of interactions among fundamental particles” (Heil 2005, p.31)

- Institutions and other entities could be introduced in the domain of quantification because this reduction cannot be expressed in the considered language or it is too complex to be (cognitively and/or computationally) managed

- *Example* 3. Points and regions

  Regions can be reduced to sets of points and points can be reduced to regions using a second order construction

  - What are the ‘real’ spatial entities? (*Quine’s ontological relativity*)

- Suppose to have evidences that points are abstractions from regions

  - Without a 2ord logic we cannot explicitly represent the construction of points from regions
  - With points and regions in the domain of quantification it is possible to conduct some analysis of their connection also in FOL

- *Example* 4. Qualities in DOLCE

  The color of the rose changed from scarlet to crimson

  - physical object: $\texttt{rose\#1}$
  - quality: $\texttt{col}_{\texttt{rose\#1}}$
  - regions: $\texttt{scarlet}, \texttt{crimson}, \texttt{red}, \texttt{colored}$

  - $\mathsf{P}(\texttt{scarlet}, \texttt{red})$, $\mathsf{P}(\texttt{crimson}, \texttt{red})$, $\mathsf{P}(\texttt{red}, \texttt{colored})$
  - $\mathsf{qt}(\texttt{col}_{\texttt{rose\#1}}, \texttt{rose\#1})$
  - $\mathsf{ql}(\texttt{scarlet}, \texttt{col}_{\texttt{rose\#1}}, \texttt{rose\#1}, \texttt{t})$, $\mathsf{ql}(\texttt{crimson}, \texttt{col}_{\texttt{rose\#1}}, \texttt{rose\#1}, \texttt{t}')$

- To conclude
  - Multidisciplinarity may help in finding solutions to representation problems
  - The choice of a formal language is delicate (analytical vs. applicative role)
  - No need for a monolithic ontology, a library of ⊚ntologies is richer and allows the user to select the view that better fits his requirements without precluding the possibility to (partial) integrate with systems based on different ontologies
- All this has a cost
  - several ⊚ntologies must be carefully developed and integrated (and the user needs to understand them to chose the most adequate for his requirements) and developers need to have a background in several disciplines

...but, as *Nicola Guarino* says

It is hard?

Of course yes!

Why should it be easy?

...but these are only slogans