How Formal Ontology can help Civil Engineers

Stefano Borgo

Laboratory for Applied Ontology, ISTC-CNR, Trento (IT)
borgo@loa-cnr.it
www.loa-cnr.it

1 Introduction

In this paper we report some considerations on the developing relationship between the area of formal ontology and that of urban development. Even in the studies on urban and territorial systems we register a phenomenon common to most applied domains: the increasing interest on ontology and the difficulties to understand its novelty. Indeed, the area of applied ontology spans a variety of methods and ideas, some of which have been developed much earlier in other approaches. This older group of 'ontological tools' (among which we find classification methods, taxonomic organization, graph and lattice theories) are well-known techniques and form the basis of most university programs (from engineering to geography, from computer science to cognitive science). It is natural that the domain experts that want to introduce applied ontology to their domain find easy to get hold of these old techniques since, in a sense, these are already part of their background. Unfortunately, these techniques have already reached their limits and now have little to say in ontology research:[1] they are substantially the same as thirty or forty years ago (even relatively recent proposals like dynamic taxonomies are just innovative applications of well-known knowledge techniques).

In contrast, it is harder for non-ontologists to understand the new ideas and techniques that applied ontology has to offer since they often are obtained by mixing ideas from disparate field like philosophy, region-based geometry and logic. This fact is not surprising because ontology is a recent and innovative area of research which has not found a proper place in education programs yet. A few compelling aspects can be identified: ontological research aims at general principles and rules which make it more abstract than the previous approaches to knowledge representation (consider the conceptual shift from the discussion of 'data' to that of 'entity' or even 'possible entity'). It applies subtle distinctions imported from the philosophical domain (like substance vs accident, tropes vs properties) which are new in conceptual modeling. Furthermore, it concentrates on good and deep formalizations of the adopted concepts (thus breaking away from the limits of conceptual systems). The combination of these and other elements explain the novelty of applied ontology and the problems it has to be properly understood by practitioners.

[1] This claim does not want to contrast their usefulness which is even higher today essentially for the improvement of modern informatics systems. They are valuable tools and are successfully applied in many situations. Nonetheless, they are of less interest (since not innovative) in ontology research as developed from the late 90s.

In what follows, we address some (and somehow scattered) issues of interest to civil engineers, architects, and experts in urban development that are sensitive, on the one hand, to the theoretical foundations of their domain area and, on the other hand, to improve the stability and reusability of their models via ontological techniques. However, before we can introduce these issues (the problem of incompatible space representations, consistent use of linguistic resources, integration of existing and disparate domain ontologies) we need to set some basic distinctions that serve us to put some order on the class of ontological systems. After all, we need to agree on what we mean by 'ontology' if we want to consistently compare alternative views and arguments.

2 Classifying Ontologies

Ontology systems (or simply ontologies) are complex systems that can be analyzed from a variety of perspectives: language, content, taxonomic structure, domain coverage, semantics and so on. Each perspective provides a different way to classify ontologies. Here it suffices to look at two of them, namely, the semantics and expressivity of the adopted language and the generality of the included concepts.

The first classification gives us a way to classify ontologies according to the language and the type of semantics it adopts. This is a crucial distinction: ontological systems are not simple classification structures, they are supposed to classify entities according to their *essential nature*. We can capture it in the ontology only through a careful use and interpretation of the adopted language. Since the major tool we have to ensure the correct interpretation of the language is formal semantics, it is important to know in which *semantic class* the ontology is positioned. Here we identify three general classes. The first includes the systems with the weakest semantics (in terms of formal semantics) since they necessarily rely on natural language. This class collects mainly linguistic and terminological ontologies, comprising the vast majority of ontologies today available. A second class includes systems usually limited to weak formal languages. The main concerns in developing ontologies in this second class are related to complexity, feasibility, and other implementation issues (which affect the generality of these systems). In the third class we find quite expressive logical theories with full formal semantics.

Once we have the semantic classes available, we can look at the formal expressivity of each system (the formal distinctions that the system can consistently make) to refine the classification. (Note that the subclasses provided here are not exhaustive.)

1. **Linguistic/Terminological ontologies**
   [these are ontologies committed primarily to the semantics of natural languages]
   - Glossary
   - Controlled vocabulary
   - Taxonomy
   - Thesaurus
2. **Implementation driven ontologies**
   [in these systems the primitives are committed to natural language semantics and the derived terms to formal semantics]

- Conceptual Schema
- Knowledge Base

3. **Formal ontologies**
   [these ontologies commit exclusively to the semantics of formal languages]
   (types are given by classes of interpreted languages like modal, predicative logics, logics with binary relations only, logics with restricted models, etc.)

The other classification instrumental to our goals is independent from the above and looks at the concepts the ontology uses to categorize entities. Such a classification is harder to provide since content is hard to define. Fortunately, for our goals it suffices to consider a rough and general classification regarding primarily the broadness of the concepts included in the systems (see also [1]).

1. Domain ontologies
2. Core (reference) ontologies
3. Foundational ontologies

2.1 Formal ontologies: the notion

Most people rely on a widely cited description of ontology which says: "An ontology is an explicit specification of a conceptualization." ([2], Sect.2). We think that the general acceptance of this notion is due in large part to the lack of constraints it puts; any collection of terms, graph of classes, and logical theory can be seen as an ontology according to the above notion. Nonetheless, Gruber's proposal gives an important intuition on what an ontology is. Then, it is important to find a technical definition that correctly separates proper ontological systems from others.

Formal ontology [3] explicates and deepens Gruber's intuition. Guarino's proposal is to add specific constraints in order to avoid misinterpretation (and misuse) of the system. In his view, an ontology must be based on:

I) a set of basic linguistic elements and a set of precise rules to construct terms and relations (adoption of a *formal* language)
II) a clearly stated semantics for the language (adoption of a *formal* semantics)
III) a rich set of explicit motivations and arguments, possibly with references to the philosophical and ontological literature, to justify and illustrate the adopted categories and relations (presence of documented *philosophical* analysis)

The above requirements constrain the technical aspects of an ontology without affecting the content. This choice makes clear that applied ontology is a scientific domain that looks at the *formal* properties of the entities it studies, i.e., the ontological systems. Regarding the content, condition III) sets a minimal request: it requires it to be well documented. No restriction is put on the view the ontology professes since this aspect is what determines its acceptance as a knowledge representation tool, not its quality as an ontological system.

With the above definition of formal ontology, it becomes possible to split the complexity of standard knowledge representation systems into two distinct parts that, by and large, correspond to the ontological component and the knowledge-base component. The first, which is the domain of formal ontologies, deals with the organization of the knowledge structure while the latter is concerned with the information contained in the knowledge structure.

2.2 An example we are all familiar with: MATH

We all have been exposed to mathematics and understand the basics. The isolation of the mathematics domain and the precision of its objects and techniques make this science suitable for challenging our intuition on what ontology is.[2] The classification of page 2 suggests that we may give several different answers. An analysis of the proposed ontologies for maths helps us since it allows u to make explicit the position we take in this paper. The reader should try to write down its own answer and compare it with the one we give below.

First, recall that mathematics is a specific language formed by terms, sentences, function symbols, quantifiers, etc. which is used to talk about special entities like sets (e.g. $\emptyset$), numbers (e.g. π), ordinals (e.g. $\aleph_0$), functions (e.g. log_e), matrices (e.g. $\left[\begin{smallmatrix} 0 & 2 \\ 3 & 3 \end{smallmatrix}\right]$) etc.[3] The entities are individuated via primitives (which come together with an axiomatization) and definitions (derived notions).

Everyone would accept that neither a language, nor a collection of entities is *per se* an ontology. This observation holds as well for the language of mathematics and the set of its entities. We continue that the collection of primitives and derived notions of mathematics (let them be concepts or relations) is tantamount not an ontology. Indeed, from the perspective embraced in this paper, we conclude that the ontology of mathematics is the *complex structure of relationships connecting primitives (as concets) and derived notions*.

2.3 What is a (formal) ontology then?

Leaving aside the variety of things people mean when using the term ‘ontological system’ or ‘ontology’ for short (a labeled graph, a set of terms, a knowledge base, a structure for knowledge etc.), one must recognize that there is a clear-cut distinction between a system *for knowledge organization* and a system *of knowledge*.

As we said, ontologies are developed to cover the first of these two senses, i.e., they are systems developed to organize knowledge. More than that, the success of the term ‘ontology’ is due, in our view, to its explicative import which is realized only when the system is coupled with a description of the view on the ‘world’ (or domain of interest) that has motivated it. Unfortunately, some researchers minimize this aspect and claim that the ontology structure itself suffices as an (implicit) description of the ontology viewpoint. Then, they do not feel committed to go further in analyzing the ontological aspects purported by the system. Most systems in the class of terminological ontologies are a consequence of this ‘permissive’ reading of the notion of ontology. Others

[2] Clearly, we posit the question from the perspective of applied ontology. The ontology of mathematics from the perspective of the philosophy of math is a different (although related) issue.

[3] Of course, in all these examples we refer to the denotations of the listed terms or expressions.

work with weak languages in which one cannot formalize even quite basic constraints. This is the source of another important fault of several systems: insufficient (actual) formalization.

We think that proper ontologies must address two main aspects:

- the *structural* aspect: the system clearly establishes and describes the types of existing entities, the structural organization and relationships among the types
- the *formal* aspect: the system is constrained with a sufficiently rich axiomatization that rules out (most) possibilities of misinterpretation

2.4 ...and what is a foundational ontology?

Foundational ontologies are formal ontologies that provide a structure for the most general types of entities. They characterize the meaning of general terms like entity, event, process, spatial and temporal location (as opposed to drilling machine, driving, being in London, the 2004 olympics) and basic relations like parthood, participation, dependence, and constitution (as opposed to mechanical parthood, playing a card game, depending on water, having an arm).

The purpose of foundational ontologyies is abstracted away from any direct application concern. These systems aim to provide a formal description of entity types and relationships that are common to *all domains* and to provide a consistent and unifying view of 'reality' from a given perspective. In principle, any (consistent) ontology is justified by a foundational ontology, i.e., by a general view on what exists and how (ontological) classes of things are related.

3 The DOLCE ontology

DOLCE [4] stands for the Descriptive Ontology for Linguistic and Cognitive Engineering. It is a foundational ontology that concentrates on *particulars*, that is, roughly speaking, objects (both physical and abstract), events, and qualities. It does not classify properties and relations: these are included in the system as far as needed to characterize particulars. DOLCE adopts the distinction between objects (like houses and refrigerators) and events (like cutting and visiting) and differentiates among individual qualities, quality types, quality spaces, and quality values as we will see. Technically, it is a formal ontology that relies on a very expressive language, first-order modal logic.

DOLCE adopts a *multiplicative approach* since it assumes that different entities can be co-localized in the same space-time. For example, a building and the amount of matter that constitutes it are captured in DOLCE as two distinct entities (as opposed to different aspects of the same entity). The reason lies on the different set of properties that these entities enjoy: the building ceases to exist if it collapses due to a earthquake since a radical change of shape occurs while the amount of matter is not affected (the identity of an amount of matter is not affected by the change of the shape). For a different example (discussed at length in the philosophical literature), consider a statue made of clay. DOLCE models the statue and the clay as different entities which share the same spatial (and possibly temporal) location. This allows us to capture the strong intuition that a scratched statue has changed (since scratched) and yet it is the same statue it was before. In DOLCE these claims are consistent since the statue itself might not be affected by (minor) scratches, but the clay (which is the constituent entity of the statue) does because amounts of matter cannot loose parts.

The category of *endurant* collects entities like a “railroad” or material like “some cement”, while events like “making a hole” and “driving a car” are in the category of *perdurant*. The term ‘object’ itself is used in the ontology to capture a notion of unity or wholeness as suggested by the partition of the class “physical endurant” into the classes “amount of matter” (whose elements are (an amount of) gold, air etc.); “feature” (a hole, a corner); and “physical objects” (a building, a human body). See Figure 1. Note that the terminology adopted departs sometimes from the usage in the knowledge representation area since it has been affected in part by philosophical literature.

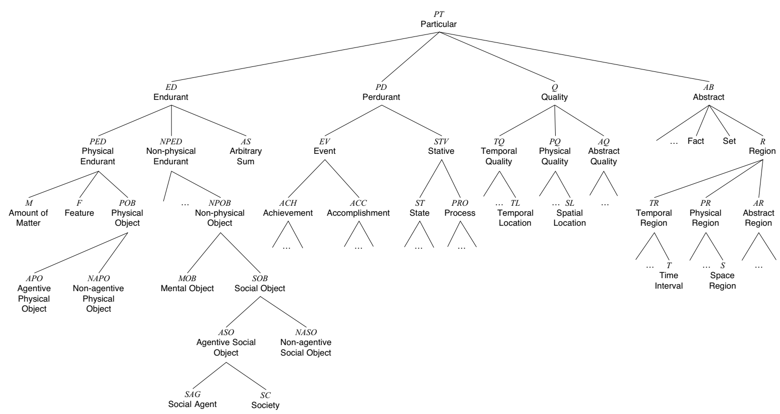

Fig. 1. Taxonomy of DOLCE basic categories (from [4]).

Both endurants and perdurants are associated with a bunch of *qualities*. These entities and their evaluation are crucial in DOLCE and the distinction between *individual qualities*, *qualia*, and *quality spaces* has been set with the aim of capturing common sense in a coherent and consistent way as we are going to see.

Qualities and Incompatible Representations

Adopting the DOLCE perspective, one can set a framework [5] where different forms of quality representations can coexist and consistently (as well as coherently) interrelate. The basic entities, as mentioned above, are *individual qualities*, e.g. the weight of this brick. Individuals qualities inhere in specific individuals so that the weight of this brick is different from the weight of that brick, no matter how similar they are. Furthermore, individual qualities can change through time since the weight of this brick matches 2 kg now and will match 1.9 kg after I cut off a corner piece. *Qualia*, e.g. a specific weight, form another type of entities. These entities are obtained by abstracting all possible individual qualities from time and from their hosts. Then, differently from individual qualities, qualia are not entity dependent. Nonetheless, analogously to individual qualities, qualia are divided in types: weight qualia, shape qualia, color qualia, and so on. If two bricks put straight the pivot of a perfect balance, then they have the same weight quale although they have different individual weight qualities. In this sense qualia represent perfect and objective similarity between (aspects of) objects. Finally, *spaces* corresponds to different ways of organizing qualia. They are motivated by subjective (context dependent, qualitative, applicative, etc.) similarity between (aspects of) objects. By means of spaces, a structure can be imposed on qualia (for example ordering, metrics, geometry and qualitative relations) and this makes it possible to differentiate several quantitative and qualitative degrees of similarity. With these distinctions it becomes possible to talk about the weight of a building in different ways as indicated by the first column of Fig. 2. Analogously, for the other qualities.

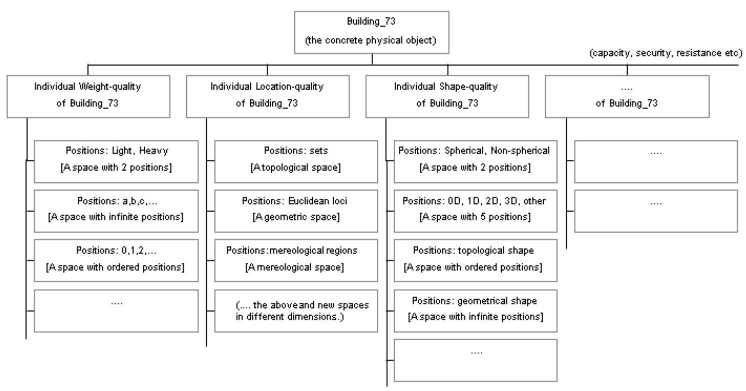

Fig. 2. Space Modularity in DOLCE.

This modularization techniques allows the use of different space representations within the same ontology. Indeed, location in DOLCE is simply an individual quality that physical entities must possess. The comparison of the location of one object with the location of another is carried out consistently in any space of interest (or even across spaces) as suggested by column 2 of Fig. 2.

4 Coupling Foundational and Weaker Ontologies

Even the optimistics would admit that it will take many years before a rich formal ontology that covers the urban development domain can be available. Also, one may doubt that such a system is needed in practice. The solution might be to find a good balance between the time- and resource-consuming effort that the development of a reliable formal ontology requires and the inexpensive and prompt availability of terminological domain ontologies. Fortunately, the adoption of a foundational ontology already suffices to greatly improve the robustness and interoperability of existing (implementation driven or even terminological) domain ontologies. From this observation, what is necessary is a careful extension of the foundational ontology with appropriate concepts that correctly organize the main categories to which core and domain concepts can be connected. This view brings forward the interesting problem of coupling foundational and weaker ontologies. The analysis of the problem (including the study of proposed solutions which in the literature are mostly based on the WordNet linguistic resource [6]) shows that different techniques can be applied.

There are basically four major strategies [7]:

1) *Re-structuring*. The ontology is used at the meta-level only. The real focus is an ontological improvement of the linguistic resource that does not require the addition of ontological categories or relations. In particular, the computational properties of the linguistic resource are unaffected.
2) *Populating*. The ontological and linguistic systems are here treated as simple taxonomies. The focus is on the mapping between these two taxonomies. The map is then used to enrich the ontology with lexical information.
3) *Aligning*. In this case the focus is on both the ontology's structure level and the linguistic object level. This approach consists in implementing both the previous perspectives of re-structuring and populating. The result, which cannot be reduced to any of the original systems, is ontologically sound and linguistically motivated.
4) *Merging*. The first step consists in isolating a system that takes the common parts of the ontology and the lexical resource. Then, the system is extended (by choosing among the alternative views given by the original systems) to ensure enough coverage. The approach relies on techniques for redundancy removal and consistency preservation.

5 Appling a Foundational Ontology

A final remark is in order: foundational ontologies *are* implementable. However, even if a foundational ontology is fully implemented, it cannot be used in the same way as terminological ontologies. The two types of systems have different roles [3] as we mentioned earlier.

The DOLCE foundational ontology is available in first-order modal logic and has several versions in different languages[4] like KIF, OWL-DL, DAML+OIL and RDFS. The *Common Algebraic Specification Language* (CASL), developed by The Common Framework Initiative [8], has been enriched with an extension, HETS, to manage foundational ontologies and their modularization; the full DOLCE ontology (including a partial modularization) is now available in the CASL system as shown in [9]. In particular, the possibility to manipulate ontologies as modular systems is crucial when dealing with large logical theories like DOLCE. Indeed, the special approach of CASL to ontology construction borrows from research in logical studies and software engineering, and is driven by applicative concerns. As a result, in a system like CASL, it becomes possible to store several domain ontologies and to reliably transfer information from one another provided they are linked to a common foundational ontology like DOLCE.

[4] For further information, visit http://www.loa-cnr.it/DOLCE.html

At the cost of complicating the system, one can even adopt different foundational ontologies, each connected to a group of domain ontologies, and transfer (part of the) available information through ontological systems that embrace very different views on 'reality'.

Acknowledgement

Most of the work reported in this paper is the result of research developed at the Laboratory for Applied Ontology (LOA) in Trento. The author has been supported by the Provincia Autonoma di Trento and the national project TOCAI.IT.

References

1. Borgo, S., Gangemi, A.: At the core of core ontologies. In Gangemi, A., Borgo, S., eds.: Proceedings of the Workshop on Core Ontologies in Ontology Engineering, CEUR-WS.org/Vol-118 (2004) 1–4
2. Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing. International Journal of Human-Computer Studies **43** (1995) 907–928
3. Guarino, N.: Formal ontology in information systems. In Guarino, N., ed.: Proceedings of the Second International Conference on Formal Ontology in Information Systems, IOS Press (1998) 3–15
4. Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A.: Ontology Library (Wonder-Web Deliverable D18). Available at http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf (2003)
5. Masolo, C., Borgo, S.: Qualities in formal ontology. In Hitzler, P., Lutz, C., Stumme, G., eds.: Foundational Aspects of Ontologies (FOnt), Universitat Koblenz (2005) 2–16
6. Fellbaum, C.e.: WordNet An Electronic Lexical Database. Bradford Book (1998)
7. Prevot, L., Borgo, S., Oltramari, A.: Interfacing ontologies and lexical resources. In: Ontologies and Lexical Resources: IJCNLP-05 Workshop. (2005) 1–12
8. CoFI: CASL Reference Manual. LNCS 2960 (IFIP Series). Springer (2004)
9. Luettich, K., Mossakowski, T.: Specification of ontologies in casl. In Varzi, A., Vieu, L., eds.: Proceedings of the Thrird International Conference FOIS 2004. Volume 114., IOS Press (2004) 140–150