In Alex Borgida, Vinay Chaudhri, Paolo Giorgini, Eric Yu (eds.), Conceptual Modelling: Foundations and Applications. Essays in Honor of John Mylopoulos, Springer Verlag 2009, pp. 52-67.

The Ontological Level: Revisiting 30 Years of Knowledge Representation

Nicola Guarino

ISTC-CNR, Laboratory for Applied Ontology, Via alla Cascata 56/C, Trento, Italy
nicola.guarino@cnr.it

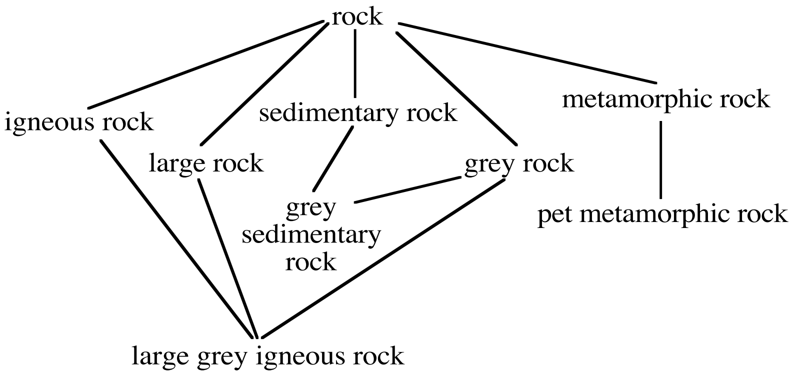
Abstract. I revisit here the motivations and the main proposal of paper I published at the 1994 Wittgenstein Symposium, entitled "The Ontological Level", in the light of the main results achieved in the latest 30 years of Knowledge Representation, since the well known "What's in a link?" paper by Bill Woods. I will argue that, despite the explosion of ontologies, many problems are still there, since there is no general agreement about having ontological distinctions built in the representation language, so that assumptions concerning the basic constructs of representation languages remain implicit in the mind of the knowledge engineer, and difficult to express and to share. I will recap the recent results concerning formal ontological distinctions among unary and binary relations, sketching a basic ontology of meta-level categories representation languages should be aware of, and I will discuss the role of such distinctions in the current practice of knowledge engineering.

Keywords: ontology, knowledge representation, identity, rigidity, OntoClean

1 Introduction

About 25 years ago, Ron Brachman, Richard Fikes and Hector Levesque [5] published a seminal paper describing a hybrid knowledge representation system (KRYPTON) built around two separate components reflecting the natural distinction between terms and sentences: the TBox (for *terminological knowledge*) and the ABox (for *assertional knowledge*). Terms were represented in the TBox by a structured formalism that was an ancestor of modern description logics, allowing the knowledge engineer to form composite descriptions corresponding to noun phrases like "an igneous rock", "a grey rock", or "a family with no children". A terminological knowledge base can be seen as a network of *analytic relationships* between such descriptions. If the basic vocabulary and the description-forming rules are rich enough, such a network can easily become quite complicated, due to the possibility of forming complex descriptions. For instance, even with a small set of attributes denoting different properties of rocks, it is easy to come up with a relatively complex taxonomy, as the authors point out while presenting Fig. 1.

In this context, the authors discussed the effects of a query such as “How many rock kinds are there?”. They observed that, despite its commonsense simplicity, this is a “dangerous question to ask”, as it cannot be answered by simply looking at the nodes subsumed by ‘rock’ in the network, since the language allows them to proliferate easily, as soon as new attributes are added to the vocabulary. Hence they proposed a functional approach to knowledge representation designed to only answer “safe” queries that are about analytical relationships between terms, and whose answers are independent of the actual structure of the knowledge base, like “a large grey igneous rock is a grey rock”.



Fig, 1. Kinds of rocks (From [5])

It is clear that, in this example, Brachman and colleagues understood the term “rock kind” in a very simple, minimalist way (perhaps as synonymous with “rock class”), ignoring the fact that, for many people, there are just three kinds of rocks, as taught at high school: Igneous, Metamorphic, and Sedimentary. On the other hand, two of the same authors, in an earlier paper on terminological competence in knowledge representation [6] stressed the importance of distinguishing an “enhancement mode transistor” (which is “a kind of transistor”) from a “pass transistor” (which is “a role a transistor plays in a larger circuit”).

So why was this distinction ignored? My own conclusion is that important issues related to the different ontological assumptions underlying our use of terms have been simply given up while striving for logical simplification and computational tractability. As a consequence, most representation languages, including “ontology languages” like OWL, do not offer constructs able to distinguish among terms having similar logical structure but different ontological implications. In our example, clearly “large rock” and “sedimentary rock” have the same logical structure, being both interpreted as the conjunction of two (primitive) logical properties; yet we tend to believe that there is something radically different between the two: why? To answer this question we have to investigate:

- the nature of the primitive properties “being a rock”, “being large”, and “being sedimentary”;
- the way they combine together in a structured term, while modifying each other.

Unfortunately, while current representation languages offer us powerful tools to build structured descriptions whose formal semantics is carefully controlled to provide efficient reasoning services, still no agreement has been reached concerning the need to adopt proper mechanisms to control the ontological commitments of structured representation formalisms, as their semantics is completely neutral with respect to the nature of the primitive components and the structuring relationships.

To see another instance of this unfortunate situation, involving binary relations instead of unary properties as in the previous case, consider the old example brought about by Bill Woods’ in its classic “What’s in a Link?” paper [38]:

```
JOHN
    HEIGHT: 6 FEET
    HIT: MARY
```

As Woods observed, in this case the two relations ‘Height’ and ‘Hit’ have certainly a different ontological nature, but nothing excludes them from being considered as “attributes” or “roles” (in the description logic’s sense), since, in the standard semantics of structured representation formalisms, these constructs are understood as *arbitrary* binary relations. So, more than 30 years later, Woods’ problem cannot be considered as solved.

Indeed, ontologies have exploded nowadays, but many problems are still there: we have now ontology languages, but despite a fair amount of results concerning the formal analysis of ontological distinctions like the ones mentioned before – including OntoClean [20, 21] and the related work on the ontological characterization of unary properties [18, 19, 31], as well as extensive analyses of fundamental binary relations such as parthood, location or dependence [32, 37, 2, 9, 33, 34, 12] – there is still no general agreement about having such distinctions built in the language, so that the assumptions concerning the basic constructs of representation languages tend to remain implicit in the mind of the knowledge engineer, and are anyway difficult to express and share. A concrete proposal in this direction has been made in [23], where an ontologically well-founded profile for UML is proposed, which constrains the semantics of UML modeling elements in the light of ontological distinctions mainly inspired to OntoClean. This is still a preliminary work, however, and we are far from having an ontologically well-founded representation language we can reason with. Moreover, nobody has explored, as far as I am aware of, the computational impact of a representation language whose semantics is constrained in the light of ontological distinctions.

In the following, I will revisit the motivations and the main proposal of my old 1994 paper [17] in the light of the main results achieved so far, arguing for the need of further work[1]. This paper is organized as follows. In the next section I will discuss the very notion of "levels" for knowledge representation languages, based on a classic paper by Ron Brachman [4], and I will argue in favor of the introduction of a specific ontological level. Then, in Section 3, I will present examples showing the practical necessity of an explicit ontological commitment for representation constructs. In section 4, I will recap the recent results concerning formal ontological distinctions among unary and binary relations, sketching a basic ontology of meta-level categories representation languages should be aware of. In section 5, I discuss the role of the ontological level in current practice of knowledge engineering.

2 Knowledge Representation Levels

Level	Primitives
Implementational	Memory cells, pointers
Logical	Propositions, predicates, functions, logical operators
Epistemological	Concept types, structuring relations
Conceptual	Conceptual relations, primitive objects and actions
Linguistic	Linguistic terms

Fig. 2. Classification of primitives used in KR formalisms (adapted from [4]). Epistemological level was "the missing level".

In 1979, Ron Brachman discussed a classification of the various primitives used by KR systems at that time [4]. He argued that they could be grouped in four levels, ranging from the implementational to the linguistic level (Fig. 2). Each level corresponds to an explicit set of primitives offered to the knowledge engineer. At the *implementational level*, primitives are merely pointers and memory cells, which allow us to construct data structures with no a priori semantics. At the *logical level*, primitives are propositions, predicates, logical functions and operators, which are given a formal semantics in terms of relations among objects in the real world. No particular assumption is made however as to the nature of such relations: classical predicate logic is a general, uniform, neutral formalism, and the user is free to adapt it to its own representation purposes. At the *conceptual level*, primitives have a definite

---

[1] Most of the material presented here has been used in PhD courses on "Foundations of Conceptual Modeling and Ontological Analysis" John Mylopoulos and I have been giving for a couple of years (with slight changes in focus) at the ICT International School of the University of Trento. The idea was to present our own approaches in a complementary way, being both present throughout the course and making comments on each other's lectures on the fly. A lot of fun.

cognitive interpretation, corresponding to language-independent concepts like elementary actions or thematic roles. Finally, primitives at the *linguistic level* are associated directly to nouns and verbs of a specific natural language.

Brachman's KL-ONE [4,7] was the first example of a formalism built around these notions. Its main contribution was to give an epistemological foundation to cognitive structures like frames and semantic networks, whose formal contradictions had been revealed in the famous "What's in a link?" paper [38]. Brachman's answer to Woods' question was that *conceptual links* should be accounted for by *epistemological links*, which represent the structural connections in our knowledge needed to justify conceptual inferences. KL-ONE focused in particular on the inferences related to the so-called IS-A relationship, offering primitives to describe the (minimal) formal structure of a concept needed to guarantee "formal inferences about the relationship (subsumption) between a concept and another". Such formal structure consisted of the basic concept's constituents (primitive concepts and role expressions) and the constraints among them, independently of any commitment as to:

- the meaning of primitive concepts;
- the meaning of roles themselves;
- the nature of each role's contribution to the meaning of a specific concept.

The intended meaning of concepts remained therefore totally arbitrary: indeed, the semantics of current descendants of KL-ONE, description logics, is such that concepts correspond to *arbitrary* monadic predicates, while roles are *arbitrary* binary relations. In other words, at the epistemological level, emphasis is more on *formal reasoning* than on (formal) *representation:* the very task of representation, i.e. the structuring of a domain, is left to the user.

Current frame-based languages and object-oriented formalisms suffer from the same problem, which is common to all epistemological-level languages. On the one hand, their advantage over purely logical languages is that some predicates, such as those corresponding to types and attributes, acquire a peculiar, *structuring* meaning. Such meaning is the result of a number of ontological commitments, often motivated by strong cognitive and linguistic reasons and ultimately dependent on the particular task being considered, which accumulate in layers starting from the very beginning of the process of developing a knowledge base [11]. On the other hand, such ontological commitments remain hidden in the knowledge engineer's mind, since these knowledge representation languages are in general *neutral* as concerns ontological choices. This is also, in a sense, a result of the *essential ontological promiscuity* claimed by influential scholars [13, 27] for AI languages: since conceptualizations are our own inventions, then we need the maximum freedom for interpreting our representations.

In my 1994 paper I argued against this neutrality, claiming that a rigorous ontological foundation for knowledge representation can improve the quality of the knowledge engineering process, making it easier to build at least understandable (if not reusable) knowledge bases. After all, even if our representations are ontologically promiscuous, admitting the existence of whatever is relevant for us, it seems certainly useful to avoid at least the most serious ontological ambiguities when it comes to interpretation, by using different constructs for different basic ontological categories. In this view, as we shall see, “being large” and “being a rock” are represented by different constructs, whose semantics is constrained to reflect general ontological distinctions.

Level	Primitive constructs	Main feature	Interpretation
Logical	Predicates	Formalisation	Arbitrary
Epistemological	Structuring relations (concepts and roles)	Structure	Arbitrary
Ontological	Structuring relations *satisfying meaning postulates*	*Meaning*	*Constrained*
Conceptual	Cognitive primitives	Conceptualisation	Subjective
Linguistic	Linguistic primitives	Language	Subjective

Fig. 3. Main features of the ontological level.

Representation languages conforming to this view belong to the *ontological level*, a new level I proposed to include in Brachman’s layered classification, in an intermediate position between the epistemological and the conceptual levels (Fig. 3). While the epistemological level is the level of *structure*, the ontological level is the level of *meaning.* At the ontological level, knowledge primitives satisfy formal meaning postulates, which restrict the interpretation of a logical theory on the basis of formal ontological distinctions.

3 From the logical level to the ontological level

Suppose we want to state that a red apple exists. At the *logical level*, it is straightforward to write down something like

(1) $\exists x\ (\mathrm{Apple}(x) \wedge \mathrm{Red}(x))$.

At the epistemological level, if we want to impose some *structure* on our domain (dividing for instance apple from pears), the simplest formalism we may resort to is many-sorted logic. Yet, we have to decide which predicates correspond to sorts, as we may write

(2) $\exists x{:}\mathrm{Apple}(\mathrm{Red}(x))$

as well as

(3) $\exists x{:}Red(Apple(x))$

or maybe

(4) $\exists(x{:}Apple, y{:}Red)(x=y)$.

All these structured formalizations are equivalent to the previous one-sorted axiom, but each contains an implicit structuring choice. However, (3) sounds intuitively odd: what are we quantifying over? Do we assume the existence of "instances of redness" that can have the property of being apples?

Unfortunately, the formalism we are using does not help us in making the right choice: we have the notion of "sort", but its semantics is completely neutral, since a sort may correspond to an *arbitrary* unary predicate. Using a more structured formalism allowing for attributes or (so-called) roles, like a description logic or a frame-based language, does not help, since we still have to make a choice between, say

(5) (a Apple with Color red)

and

(6) (a Red with Shape apple)

So, at the epistemological level, the structuring choices are up to the user, and there is no way to exclude the "unnatural" ones.

At the *ontological level,* on the contrary, what we want is a formal, restricted semantic account that reflects the ontological commitment underlying each structuring primitive, so that the association between a logical predicate and a structuring primitive is not a neutral choice any more: in other words, each structuring primitive corresponds to properties (or relations) *of a certain kind*. In our example, the difference between "being an apple" and "being red" lies in the fact that the former property supplies a principle for distinguishing and tracing in time its individual instances, while the latter does not. This distinction is known in the philosophical literature as the distinction between *sortal* and non-sortal (or *characterising*) properties [14], and is (roughly) reflected in natural language by the fact that the former are denoted by common nouns, while the latter by adjectives. The bottom line is that not all properties are the same, and only sortal properties correspond to what are usually called "concepts".

In the light of the above criteria, a predicate like *Red* – under its ordinary meaning – will not satisfy the conditions for being a concept (or a sort). Notice however that this may be simply a matter of point of view: at the ontological level, it is still the user who decides which conditions reflect the *intended* use of the *Red* predicate. For example, consider a different scenario for our example. Suppose there is a painter, who has a palette where the various colors are labeled with terms evoking natural things. For her, the various shades of red in the palette are labeled "orange red", "cherry red", "strawberry red", "apple red". In this scenario, the formula (3) above makes perfect sense, meaning that, among the various reds, there is also the apple red.

How can we account for such semantic differences? We shall see in the following that they are not simply related to the fact that the argument of *Red* belongs to different domains, but they reflect different *ways of predication,* expressed by predicates belonging to different *kinds*, in virtue of their different ontological nature. In part, these differences are also revealed by the way we use the same word in natural language: for instance, in the first scenario *Red* is an adjective, while in the painter's scenario it is a noun. Unfortunately this basic difference disappears when we move from linguistic analysis to logic analysis, since we tend to use the same predicate for the two cases.

In a knowledge representation formalism, we are constantly using natural language words within our formulas, relying on them to make our statements readable and to convey meanings we have not explicitly stated: however, since words are ambiguous in natural language, when these words become predicate symbols it may be important to "tag" them with an ontological category, endowed with a suitable axiomatization, in order to make sure the proper intended meaning is conveyed, and to exclude at least the most serious misunderstandings. This is basically what Chris Welty and I have suggested with our OntoClean methodology [21]. However, with my *ontological level* proposal, I was aiming at something more: embed some basic ontological categories in a knowledge representation formalism, constraining its own representation primitives. In part, this is what has been attempted by Giancarlo Guizzardi in his PhD work [24]. However, this work only concern semantic constraints on a conceptual modeling language (UML V2.0), and I am not aware of similar attempts for constraining the semantics of knowledge representation formalisms such as description logics.

In the following, I will briefly sum up and revisit the most relevant distinctions within unary properties and binary relations which have emerged from the research on formal ontology since the time I published my 1994 paper, and which I believe make sense from the point of view of knowledge representation. Hopefully, such distinctions will inspire a future generation of ontological level representation languages.

4 Basic distinctions among properties

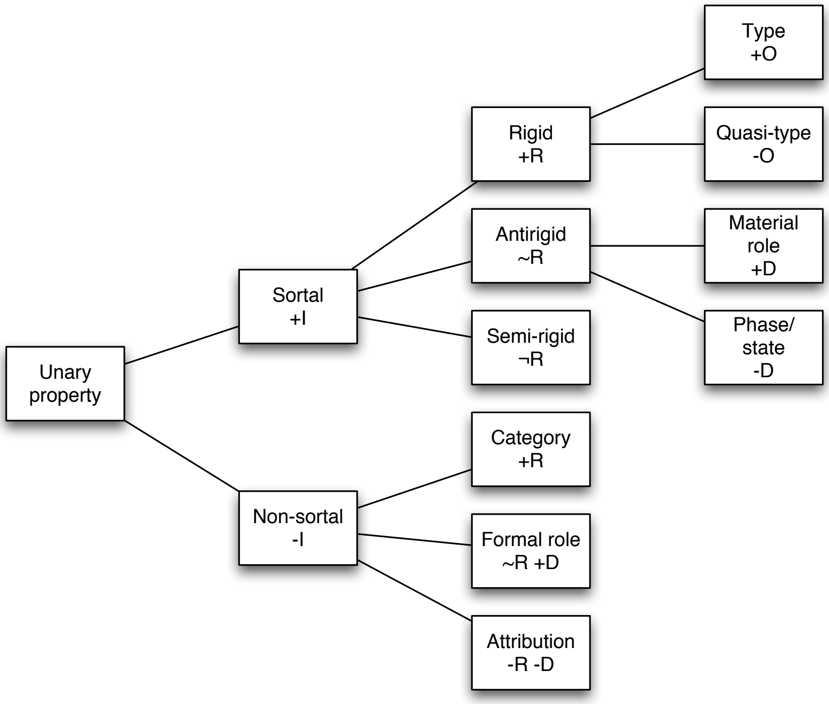


Fig. 4. A general ontology of unary properties. Adapted from [19]

In [19], Chris Welty and I presented a general ontology of unary properties, resulting from the combinatorial composition of a small set of *formal metaproperties* based on three main notions: *identity, rigidity* and *dependence*, reported (in slightly revised form) in Fig. 4. I will not go here into the details of the technical aspects underlying these metaproperties, whose formal definitions have been discussed and refined in various papers since my early proposals [16, 17, 36, 31, 24]. I will just introduce them in an informal way as needed, pointing to the most recent formalizations.

What I would like to insist on here is the practical relevance of these distinctions: not all unary properties play the same role in knowledge representation, despite the fact that all of them can be expressed by the same logical structure (unary predicate).

Before introducing these property kinds, let me stress that they are *completely general*, being independent of any commitment concerning the ontological nature of the property arguments. In other words, the reason why a certain property belongs to one of these kinds has nothing to do with its arguments, which may belong – for instance – to any of the DOLCE's top categories like objects, events, or qualities.

4.1 Sortal vs. Non-sortal Properties

The first basic distinction is the classic one between *sortal* and *non-sortal* properties. In short, a property is a sortal (marked with the meta-property +I) if it carries a *criterion of identity* for its instances. Otherwise it is a non-sortal, marked with –I.

I will not enter here in the (still well alive, see [14]) philosophical debate related to the nature of sortals, simply claiming that, especially for knowledge representation purposes, it is *extremely useful* to distinguish between properties for which a certain principle for distinguishing and tracing their instances can be determined, and properties for which such principle cannot be determined[2]. Indeed, besides being well recognized in philosophy and in linguistics, the role of identity principles is explicitly defended in conceptual modeling (for instance, in Chen's Entity-Relationship model [10], entities are explicitly defined as "'things' which can be distinctly identified").

I only note here that, differently from [17] and [23] (but consistently with the OntoClean literature) I include non-countable properties corresponding to so-called *mass-terms* (like "amount of gold") under sortals. The rationale for this is that amounts of matter *can* indeed be distinguished and traced in time, differently from non-sortal properties like "red" (in the adjectival sense), and can appear in relative clauses instantiating the pattern "the X that …", such as "the amount of water that was in the glass is now on the floor". Indeed, assuming an atomic view of amounts of matter, their identity criterion is very simple: two amounts of matter are the same if and only if they contain the same molecules (similarly to collectives like groups of people). After all, we *need* to distinguish and trace amounts of matter if we want to model flow of liquids, for instance.

So being a sortal does not imply being countable, although the converse is true, at least for ordinary domains[3], and indeed countability is a useful heuristic to conclude that a property is a sortal, independently whether a particular identity criteria can be determined.

4.2 Kinds of Rigidity

I introduced the first time the notion of ontological rigidity for a unary property in [16][4]. Since then, Chris Welty and I proposed a more careful definition in our Ontoclean papers [20, 21], which was further refined by various contributions [28, 8, 1, 36]. The basic intuition is however still the same: a unary property is rigid if it is essential for all its instances, so that, if x is an instance of a rigid property, it cannot lose this property without losing its identity. Going back to our example, it seems plausible to assume that *Apple* is always rigid (+R), while *Red* is non-rigid (-R) in the first scenario, and rigid in the painter's scenario. We see therefore how clarifying

---

[2] See [20] for a formal account of the notion of identity criteria in knowledge representation.

[3] See [29] for an argument against the fact that countability implies identity

[4] I was unfortunately unaware of the work by Gupta [25], subsequently cited in [23] who introduced a very similar notion, called *modal constancy.*

whether a property is rigid or not helps disambiguating between different ontological assumptions concerning the use of a certain word.

Since the definition of rigidity involves a universal quantification on all the instances of a given property, we can isolate two forms of rigidity: in the weaker case (non-rigidity, -R) there is at least one *contingent instance*, which does not exhibit the given property necessarily; in the stronger case (anti-rigidity, ~R), all instances are contingent. Of course anti-rigidity implies non-rigidity; a property which is non-rigid but not anti-rigid is called *semi-rigid* (¬R). As we shall see, *Student* is a classic example of an anti-rigid property (since every student is not necessarily such), while *Red* can be considered as semi-rigid, if we assume that certain things (say, rubies) are necessarily red, while others (e.g., red cars) are just contingently so. As shown in Fig. 4, sortals can be partitioned in rigid, anti-rigid and semi-rigid.

As stressed many times in the OntoClean papers, I would like to remark here that, in a certain KR theory, the decision as to whether a certain property is rigid or not is not a fixed one, and ultimately depends on the knowledge engineer: for example, if one believes in reincarnation, perhaps it makes sense to assume that *Person* is not rigid, if the worlds concerning the other lives are part of the modeling context. In a recent paper addressing again the definition of an ontology [22], I have elaborated this issue suggesting that a world is defined with respect to a specific observer (the knowledge engineer) and (forgetting time for the sake of simplicity) coincides with a maximal “perception state”. So, for the knowledge engineering practice, rigidity only concerns those worlds that are in the modeler’s radar.

4.3 Rigid Sortals: Types and Quasi-types

Rigid sortals are particularly important in knowledge engineering, since they capture the *essential, invariant aspects* of individuals, providing at the same time the criteria for individuating them in a given world, and tracing them across worlds. It seems very natural therefore, as introduced in [20] and further elaborated in [24], to impose, as modeling constraint, that *every element of the domain of discourse must be an instance of a rigid sortal*, complying to Quine’s ditto “no entity without identity”. Assuming this constraint, while analyzing a domain we can concentrate first on such rigid properties, forgetting the non-rigid ones, being assured that no domain elements are left out.

Since rigid sortals can specialize each other, it is also useful to distinguish, within a sortals taxonomy, between those which just *carry* some identity criteria (inherited from some more general sortal) and those that directly *supply* the (necessary or sufficient) conditions that contribute to such criteria. We call the latter *types*, and the former *quasi-types*. According to the OntoClean notation, types are marked with the metaproperty +O, which stands for “supplies its own identity”, and quasi-types with the metaproperty –O. For instance, consider the properties *Living Being*, *Person*, and *Italian Person.* Assuming that all of them are rigid, *Living Being* supplies some identity criteria (say, DNA identity), which are further specialized by *Person*, which adds, e.g, identity of fingerprints as a sufficient condition. Presumably, *Italian Person* does not supply further identity conditions, so the former two properties are types, while the latter is a quasi-type.

4.4 Anti-rigid Sortals: Material Roles and Phases

Since the early KL-One, the notion of role has been extensively discussed in the KR literature (see [3] for a recent overview). Various issues are still open, but there is a substantial agreement on the fact that unary properties denoting roles are anti-rigid. Anti-rigidity alone is however not enough to capture the relational nature of roles, which has been called *foundation* in [16], *external dependence* in [19], and again *foundation* in [31], always with slightly different formalizations. The latter formalization (which in turn relies on the notion of *definitional dependence*) is definitely the most accurate for our purposes, but I prefer to call it again *external dependence*, just because I find the term more intuitive. So, according to this revised definition, a property P is *externally dependent* (marked with +D) if its definition involves (at least) another property Q such that, for every instance x of P, there exists an instance y of Q which is *external* to x, in the sense that x is not a part of y, and y is not a part of x[5].

In conclusion, roles are anti-rigid, externally dependent unary properties[6]. Being anti-rigid, roles do not supply any identity criteria, which in most cases are inherited by the types they specialize (as in the prototypical example *Student*, which inherits the identity criteria of *Person*). However, there are certain general roles, like *Part*, or so-called *thematic roles* like *Patient* or *Theme,* which are not conceivably subsumed by any sortal, and hence they are not sortal themselves. Within roles, we distinguish therefore *material roles*, which (indirectly) carry some identity criteria (+I) from *formal roles*, which do not carry identity (-I).

Note that within material roles we also include properties like *Pedestrian* or *By-pass capacitor,* which linguistically behave differently from *Student* or *Son.* In [16] I called the latter *relational roles*, and the former *non-relational roles* (see next section).

As we have seen, roles are *externally dependent* properties, characterized by the +D metaproperty. If such metaproperty does not hold, and still we have an anti-rigid sortal, this is a case of a *phasal sortal,* whose prototypical example is *Baby*: if somebody is a baby, we cannot assume that anything else must necessarily exist, so *Baby* is not externally dependent, while clearly being an anti-rigid sortal. Note that phasal sortals also include *states* like *Tired* or *Happy*, assuming it is a sortal inheriting

---

[5] See [31] for the formal definition, which is based on a reification on the properties P and Q. See also [35] for a general discussion on this property reification move.

[6] See below for their systematic link to *binary* properties (so that *Student* is systematically linked with *Has-Student* or *Student-of*)

identity criteria from, e.g., *Animal.* The difference between phases and states should be however further analyzed[7].

4.5 Semi-rigid Sortals

Semi-rigid sortals have been called "mixins" in our OntoClean papers, but I prefer to avoid this term since it is used with different meanings in the object-oriented literature, as discussed in [23]. I don't think semi-rigid sortals have a special role in knowledge representation, although in some cases they may correspond to useful generalizations. They are reported here just for completeness.

4.6 Non-sortals: Categories, Formal Roles, and Attributions

The bottom part of Fig. 4 describes the remaining three cases in our taxonomy of unary properties, concerning the relevant distinctions within non-sortals. Note that our assumption that every individual must be an instance of a sortal implies that non-sortals correspond to abstract classes in the UML terminology, that is, they cannot have direct instances.

A first case is that of so-called *categories*, consisting of general properties like *Entity* or *Object*, which do not exhibit any *common* criterion of identity for their instances (for this reason they have been called *dispersive* in [26]). These are usually the topmost concepts in an ontology.

Formal roles have already been discussed, they are anti-rigid and externally dependent, but they carry no identity criteria. Note that also relational properties like *Interesting*, *Strange* or *On-the-table* fit under this class, although they don't look like roles, probably because they are not denoted by a name.

Finally, in OntoClean we called *attributions* all those non-sortal properties which are simply non-rigid and not externally dependent. This is a large class, which includes *Red* and *Big* as well as *Broken*. In DOLCE, I assume that these attributions reflect *qualitative states* of entities, resulting from the fact that a specific quality is classified in a certain region of a quality space [30].

4.7 The Rocks Example Revisited

Going back to our introductory examples, it is easy to conclude, in the light of the above discussion, that *Metamorphic rock, Igneous rock* and *Sedimentary rock* are the only *types* in the picture (we might want to call them *kinds*, terminological distinctions are a matter of taste, here). *Large rock* and *Grey rock* are semi-rigid sortals or perhaps phasal sortals (depending whether we admit that the same rock can change size or color), while *Pet metamorphic rock* is a material role.

---

[7] Perhaps phases – together with material roles – supply *local* identity criteria, differently from states.

5 Basic distinctions among binary properties

Analogously to unary properties, useful distinctions can be drawn within binary properties, with the purpose of developing more "ontology aware" representation formalisms. Unfortunately, the results in this area, in comparison to what has been done for unary properties, are much more scattered, and I am not aware of any attempt to propose a general ontology like the one described above[8].

The main practical problem of binary relations, from the KR point of view, is still the one raised by Bill Woods in the example I mentioned in the introduction: how to distinguish between the relations which contribute to the *internal structure* of a concept and those which do not? Or, in other words, how to decide whether a piece of information should be modeled in terms of an attribute-value pair or in terms of a genuine relation?

I discussed this issue in [16], suggesting that attributes should be confined to *relational roles*, *qualities*, and *parts.* Intuitively, all these cases fit under the linguistic test suggested by Woods to check whether a binary relation A can be considered as an attribute for an individual X:

Y is a value of the attribute A of X if we can say that
Y is an A of X (or Y *is the A of X*)

Retrospectively, in the light of the most recent (yet scattered) work on the ontology of relations, I believe that the intuition behind the use of the *of* preposition to capture the notion of attribute lies in the ontological distinction between *internal* and *external* relations, which is intertwined with the distinction between *formal* and *material* relations[9]. The picture I have mind for binary relations is sketched in Fig. 5. I assume first a distinction between *formal* and *material relations* [15], where a formal relation yields just because of the very existence of its relata, while a material relation needs, so to speak, another "grounding" entity. Suppose, for example that John is older than Mary and John loves Mary; the *Older-than* relationship is a formal one, while the *Loves* relationship is a material one, since – besides the existence of John and Mary – it requires an extra entity, namely the *event* consisting of the love between John and Mary. I assume that all material relations are grounded on events, in DOLCE's sense[10].

---

[8] See [15] for a recent philosophical exploration of the ontology of binary relations.

[9] I know that for some authors these terminologies are equivalent.

[10] I know that this assumption may be too strong in some cases (e.g., for certain relations between events), but I believe it is robust enough for knowledge engineering purposes.

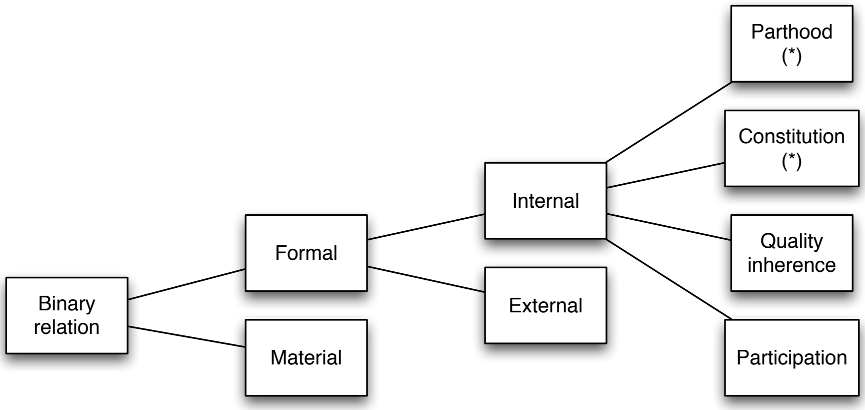


Fig. 5. A sketch of basic distinctions within binary relations

Within formal relations, I distinguish between the internal and the external ones, depending whether there is an *existential dependence* relationship between the relata. The basic kinds of internal relationships I have in mind (all formalized in DOLCE) are parthood, constitution, quality inherence, and participation, shown in the figure. There are however some technical problems concerning parthood and constitution (which are shown with an asterisk), since, if we take time into account, a specific parthood or constitution relationship can be understood as an internal relation only if it holds *necessarily* (concerning therefore an essential part); otherwise, we cannot simply say that such relationship holds without specifying the time frame (i.e. the *event*) where this happens. I don't think that explicitly modeling events involving contingent parthood or constitution is a practical choice, however, so probably the best thing is to introduce suitable time-indexed parthood and constitution relations, whose formal characterization is still being investigated. However, my suggestion in the light of this analysis is that, in an ontologically well-founded theory, structuring relations (i.e., those corresponding to are called attributes or roles in frame-based formalisms and description logics) should be limited to specializations of such internal relationships, possibly extended with time indexes. This means that, for instance, an ownership relationship between a person and her car should be modeled in terms of the entity that grounds it, namely an event to which the person and the car participate. Similarly for the *Home Address* relation, which can be expressed in terms of the location of a *Dwelling* event.

In turn, such events can be modeled in terms of their own internal relations, including the various *participation relations* (thematic relations) expressing the various ways an object participates in an event. This systematic introduction of events in place of material relations may in some cases be excessively cumbersome, but in my opinion it is the only strategy that guarantees an explicit account of the modeler's ontological assumptions. Of course, if needed, more agile relations, such as ownership, can be *defined* in terms of this more basic picture.

6 Conclusions

I hope to have shown in this paper that in order to capture the *desiderata* for knowledge representation formalisms, as expressed in the old days and never properly met, it is necessary to formally express the ontological commitments of our representation constructs. This can be done in two ways:

1. by developing general ontologies built using ontologically neutral representation constructs,
2. by adopting *non-neutral* constructs, whose semantics is suitably constrained in order to guarantee ontologically well-founded models.

I believe that the second option is preferable, since it gives the knowledge engineer the tools to produce models with certain "guaranteed" properties in terms of ontological transparency, well-foundedness, and – therefore – reusability. In addition, I believe that reasoning with such constructs should be somewhat easier than with the first option, since the expressivity required to account for their ontological commitment belongs to the meta-language (i.e., the language used to account for the ontological semantics), and not to the object language. This is however an issue to be further investigated.

References

1. Andersen, W. and Menzel, C. 2004. Modal Rigidity in the OntoClean Methodology. In L. Vieu and A. Varzi (eds.), Formal Ontology and Information Systems: Collected Papers from the Fifth International Conference. IOS Press: 119-127.
2. Artale, A., Franconi, E., Guarino, N., and Pazzi, L. 1996. Part-Whole Relations in Object-Centered Systems: an Overview. Data & Knowledge Engineering, 20(3): 347-383.
3. Boella, G., van der Torre, L., and Verhagen, H. 2007. Roles, an Interdisciplinary Perspective. Applied Ontology, 2(2): 81-88.
4. Brachman, R. J. 1979. "On the Epistemological Status of Semantic Networks". In N. V. Findler (Ed.), Associative Networks: Representation and Use of Knowledge by Computers. Academic Press, 1979.
5. Brachman, R., Fikes, R., and Levesque, H. 1983. Krypton: A Functional Approach to Knowledge Representation. IEEE Computer, 116(10): 67-73.
6. Brachman, R. and Levesque, H. 1982. Competence in Knowledge Representation. In Proceedings of National Conference on Artificial Intelligence (AAAI 82). Pittsburgh, American Association for Artificial Intelligence: 189-192.

7. Brachman, R. J. and J. G. Schmolze 1985. "An Overview of the KL-ONE Knowledge Representation System", Cognitive Science 9, 171-216.
8. Carrara, M., Giaretta, P., Morato, V., Soavi, M., and Spolaore, G. 2004. Identity and Modality in OntoClean. In L. Vieu and A. Varzi (eds.), Formal Ontology and Information Systems: Collected Papers from the Fifth International Conference. IOS Press: 128-139.
9. Casati, R. and Varzi, A. 1999. Parts and Places. The Structure of Spatial Representation. MIT Press, Cambridge, MA.
10. Chen, P. 1976. The entity-relationship model: Towards a unified view of data. ACM Transactions on Database Systems, 1(1).
11. Davis, R., H. Shrobe, et al. 1993. "What is in a Knowledge Representation?", *AI Magazine* (Spring 1993).
12. Fine, K. 1995. Ontological Dependence. Proceedings of the Aristotelian Society, 95: 269-90.
13. Genesereth, M. R. and N. J. Nilsson 1987. Logical Foundation of Artificial Intelligence. Los Altos, California: Morgan Kaufmann.
14. Grandy 2007. Sortals. In E. N. Zalta (ed.) The Stanford Encyclopedia of Philosophy (2007 edition).
15. Grenon, P. 2007. On Relations. PhD Dissertation, Dept. of Philosophy, University of Geneve
16. Guarino, N. 1992. "Concepts, Attributes and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases", Data and Knowledge Engineering 8, 249-261.
17. Guarino, N. 1994. The Ontological Level. In R. Casati, B. Smith and G. White (eds.), Philosophy and the Cognitive Science. Hölder-Pichler-Tempsky, Vienna: 443-456.
18. Guarino, N., Carrara, M., and Giaretta, P. 1994. An Ontology of Meta-Level Categories. In D. J., E. Sandewall and P. Torasso (eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Fourth International Conference (KR94). Morgan Kaufmann, San Mateo, CA: 270-280.
19. Guarino, N. and Welty, C. 2000. A Formal Ontology of Properties. In R. Dieng and O. Corby (eds.), Knowledge Engineering and Knowledge Management: Methods, Models and Tools. 12th International Conference, EKAW2000. Springer Verlag, France: 97-112.
20. Guarino, N. and Welty, C. 2002a. Identity and subsumption. In R. Green, C. Bean and S. Myaeng (eds.), The Semantics of Relationships: an Interdisciplinary Perspective. Kluwer: 111-126.
21. Guarino, N. and Welty, C. 2002b. Evaluating Ontological Decisions with OntoClean. Communications of the ACM, 45(2): 61-65.
22. Guarino, N., Oberle, D., and Staab, S. 2009. What is an Ontology? In S. Staab and R. Studer (eds.), Handbook on Ontologies, Second Edition. Springer Verlag: 1-17.
23. Guizzardi, G., Wagner, G., Guarino, N., and van Sinderen, M. 2004. An Ontologically Well-Founded Profile for UML Conceptual Models. In A. Persson and J. Stirna (eds.), Advanced Information Systems Engineering, Proceedings of 16th Intl. Conf. CAISE 04. Springer Verlag, Berlin: 112-126.
24. Guizzardi, G. 2005. Ontological Foundations for Structural Conceptual Models. Telematica Instituut Fundamental Research Series, 15.

25. Gupta, A. 1980. The Logic of Common Nouns: An Investigation in Quantified Modal Logic. Yale University Press, New Haven.
26. Hirsch, E. 1993. Dividing Reality. Oxford University Press, New York.
27. Hobbs, J. R. 1985. Ontological Promiscuity. In Proceedings of 23rd Annual Meeting of the Association for Computational Linguistics (ACL-85). Chicago, IL: 61-69.
28. Kaplan, A. 2001. Towards a Consistent Logical Framework for Ontological Analysis. In C. Welty and B. Smith (eds.), Proceedings of FOIS 2001. IOS Press: 244-255.
29. Lowe, E. J. 1998. Entity, Identity, and Unity. Erkenntnis, 48: 191-208.
30. Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A., and Schneider, L. 2002. The WonderWeb Library of Foundational Ontologies and the DOLCE ontology. WonderWeb Deliverable D17.
31. Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., and Guarino, N. 2004. Social Roles and their Descriptions. In Proceedings of 9th Intl. Conference on the Principles of Knowledge Representation and Reasoning (KR 2004). Whistler (Canada): 267-277.
32. Simons, P. 1987. Parts: a Study in Ontology. Oxford: Clarendon Press.
33. Varzi, A. 2006. A note on the transitivity of parthood. Applied Ontology, 1(2): 141-146.
34. Vieu, L. 2006. On the transitivity of functional parthood. Applied Ontology, 1(2): 147-155.
35. Vieu, L., Borgo, S., and Masolo, C. 2008. Artefacts and Roles: Modelling Strategies in a Multiplicative Ontology. In Proceedings of FOIS 2008. Saarbruecken (Germany), IOS Press.
36. Welty, C. and Andersen, W. 2005. Towards OntoClean 2.0: A framework for rigidity. Applied Ontology, 1(1): 107-116.
37. Winston, M., Chaffin, R., and Herrmann, D. 1987. A Taxonomy of Part-Whole Relations. Cognitive Science, 11: 417-444.
38. Woods, W. A. 1975. “What’s in a Link: Foundations for Semantic Networks.”. In D. G. Bobrow and A. M. Collins (Ed.), Representation and Understanding: Studies in Cognitive Science. Academic Press, 1975.