Paper to appear on *Proc. of IJCAI 2001 workshop on Ontologies and Information Sharing*

Understanding top-level ontological distinctions

Aldo Gangemi[(1)], Nicola Guarino[(2)], Claudio Masolo[(3)], Alessandro Oltramari[(2)]

1. ITBM-CNR, Rome, Italy
2. LADSEB-CNR, Padova, Italy

(3) Dept. of Electronics and Computer Science, University of Padova, Italy
nicola.guarino@ladseb.pd.cnr.it

Abstract

The main goal of this paper is to present a systematic methodology for selecting general ontological categories to be used for multiple practical purposes. After a brief overview of our basic assumptions concerning the way a useful top-level ontology should be linked to language and cognition, we present a set of primitive relations that we believe play a foundational role. On the basis of these relations, we define a few formal properties, which combined together help to understand and clarify the nature of many common ontological distinctions.

1 Introduction

1.1 Goals of this paper

The main goal of this paper is to present a systematic methodology for selecting general ontological categories to be used for multiple practical purposes. After a brief overview of our basic assumptions concerning the way a useful top-level ontology should be linked to language and cognition, we present a set of formal (i.e., domain-neutral) primitive relations that we believe play a foundational role. On the basis of these relations, we define a few further properties, which combined together help to understand and clarify the nature of many common ontological distinctions.

Our attempt is to avoid strong ontological commitments in the early steps of the methodology, trying first to establish the formal framework needed to understand, compare, and evaluate the ontological choices that ultimately will be taken.

1.2 Limits of this paper

We are conscious that our task is very ambitious, as it necessarily faces deep and highly debated philosophical and technical problems. So we have tried to be as humble as possible, making drastic simplifications whenever possible, but trying however to save the logical rigor.

One of the most serious simplifications we have made concerns the treatment of time, which is not addressed explicitly. This is in part because we believe that ontological choices about time need to be taken after a more general ontological framework is established and in part just because temporal issues are hard.

2 Ontology, cognition and language

Is ontology about the “real world” (as seen, say, by a physicist)? Or, rather, should it take cognition into account, including the complex interactions and dependencies between our ecological niche and us? We will not attempt a general answer to this question, but we believe that the latter position is very useful when building ontologies for practical purposes. As knowledge systems manage information relevant to human agents, their ontologies need to make room for entities that depend on our perception and language, and ultimately on the way we use reality to survive and communicate. Some of these entities will depend on specific groups of human beings (with their own culture, language, and so on); others will reflect common cognitive structures that depend on our sensorial interaction with reality. A general-purpose ontology is specially interested to the latter kind of entities, which help generalize our specific knowledge of the world. This position reflects the so-called “interactionist” paradigm, which (though not prevalent) has strong support in psychology of perception and cognitive linguistics [Gibson 1977, Lakoff and Johnson 1999] and seems to be a good compromise between hard 'referentialist' ontology and purely context-oriented semiotics.

An extreme example of how ontologically relevant entities depend on our perceptive and cognitive structures is the notion of *constellation*: is a constellation a genuine thing, different from a collection of stars? This is not so clear at a first sight. But, if we distinguish between stars and their specific arrangements, we are able to understand how constellations may be considered as cognitive *things* dependent on states of mind. To see a "Taurus" in the sky does not mean, obviously, that an animal is flying in the space or (less obviously) that a bull-shaped astronomic object (different from a collection of stars) is localized in a region of the sky. Rather, the perspective we embrace consists in recognizing a *cognitive entity* dependent on the way we perceive some particular arrangement of stars. It is to this entity that we (often) refer when we use the term "Taurus constellation" in our language. Including cognitive entities in our ontology seems therefore a good idea if natural language plays a relevant role in our applications.

For these reasons, we believe that a very useful and important requirement for top-level ontologies is the possibility of mapping them into large lexical databases (as, for example, WordNet [Fellbaum, 1998]). Although these large lexicons present many problems and limitations, they provide i) a source for distinctions used by humans as cognitive agents; ii) a way to give understandable names to ontological entities; iii) a practical hook towards NLP applications.

But, how do lexicons and ontologies link to each other? Assuming all lexical concepts as distinct (as for example the WordNet's *synsets*), ontologies that want to have the same conceptual coverage have to contain at least every concept of a lexicon. Adding non-lexicalized nodes to ontologies would be useful under different points of view. First of all, their presence may result in a better taxonomic organization; second, they may simplify the alignment with other ontologies and lexical sources, isolating the differences and the integration problems. So, an important (though idealistic) requirement to be satisfied would be that each term of a lexicon has an unique correspondent category in the ontology and that each ontological concept maps into at most one lexical concepts.

3 Methodology and basic assumptions

The requirement of a link with language and cognition makes the task of designing a good top-level ontology even more complicated. We outline here the methodology we suggest to accomplish such a task.

3.1 The role of formal relations

In philosophy we find a distinction between *formal ontology* and *material ontology*. Intuitively, this distinction seems to deal with the "level of generality" of ontological properties and relations, but its logical implications are not very clear. Smith gives the following "definition":

> "As formal logic deals with properties of inferences which are formal in the sense they apply to inferences in virtue of their form alone, so formal ontology deals with properties of objects which are formal in the sense that they can be exemplified, in principle, by objects in all material spheres or domains of reality." [Smith, 1998]

In this sense, we can consider *formal relations* as relations involving entities in all "material spheres", so that they are understandable *per se* as a universal notions. On the contrary, *material relations* are specific to one or more material spheres. This account seems however to presuppose an *a priori* division of the domain into "material spheres": first we establish a set of primitive subdomains (categories?), and then we distinguish between formal and material relations on the basis of their scope's behavior with respect to these subdomains. So, formal relations establish the connections and the differences between primitive subdomains; while material relations characterize, in a more detailed way, the properties of a specific subdomain. If we assume a flat domain, with no *a priori* structure, then the proposed distinction between formal and material relations collapses.

In both cases, choosing the right primitives is not easy. In one case, we must answer the question: "Which are the primitive subdomains?". In the other case, the question is: "Which are the primitive relations?". The two questions look indeed quite similar.

In this work we prefer to start with a set of primitive relations defined on a flat domain, and use them to *reconstruct* the classic categorial distinctions. This choice is mainly a matter of methodological clarity and economy, and is also motivated by our desire to maintain ontological neutrality as much as possible. These primitive relations will be still called "formal", as they will be selected among those considered as "formal" in the philosophical literature. By means of these formal relations we shall be able to:

- Formulate general constraints (e.g., atomicity) on all domain entities;
- Induce distinctions between entities (e.g., dependent vs. independent), and impose a general structure on the domain.

3.2 The methodology in a nutshell[1]

The methodology we have adopted can be summed up as follows:

1. Select from the classical philosophical repertoire a set of *formal relations* (neutral with respect to the domain choice) which shall play a *foundational role* in our ontology.
2. Select and adapt from the literature the *ground* axioms for these relations, such as those concerning their algebraic properties.
3. Add *non-ground* axioms, which establish constraints across basic relations.
4. Define a set of *formal properties* induced by the formal relations.
5. Analyze systematically the allowed combinations of formal properties, introducing a set of *basic categories*
6. Classify the relevant kinds of domain entities according to the basic categories. The result will help to understand the minimal domain structure.
7. Study the dependencies/interrelationships among basic categories, introducing *intercategorial relations.*
8. Increase the depth level of ontological analysis, by iterating this methodology within each basic category.

This work is still in progress, and in this paper we discuss in detail only points 1-4 of the methodology. However, we hope in the possibility of a progressive methodological refinement and adjustment in the way we have outlined. Moreover, we must make clear that, to start the above process, we need first some minimal assumptions (or at least intuitions) about our *largest* domain of interest (which depend on the choices discussed in section 2). We also need to

[1] See ([Thomasson, 1999] p.111-134) for an interesting discussion which advocates a methodology for ontological analysis very similar to the present one.

make some preliminary choices concerning the formal treatment of existence, modality, space and time. We shall not discuss these issues here, although we believe that these choices can be better understood, refined, or modified, by applying the methodology above.

4 Formal Relations

4.1 Instantiation and Membership

In the ontological engineering community, classical first order logic with equality is generally adopted as a formalization language (more or less reduced in its expressivity if computational efficiency is important). This means that we take for granted the distinction between properties and domain entities: the latter (syntactically denoted by constants) are usually called *instances* of the formers (syntactically denoted by predicates). The instantiation relation seems to have therefore an intrinsic meta-logical nature, as it links together entities belonging to different logical levels. Things are complicated by the fact that, given a theory A, we can construct a meta-level theory B whose constant symbols correspond to A's predicates, and whose intended domain is that of A's properties. So the term "instance" is ontologically ambiguous, unless the corresponding level is specified. There is however a bottom level, that of *ultimate instances*, things that cannot be predicated of anything else. These are what philosophers call 'particulars', i.e., entities that cannot be instantiated, as opposed to 'universals', i.e. entities that can be predicated on particulars[2].

Despite its apparent simplicity, the notion of instantiation is subtle, and should not be confused with that of set *membership*. Let's try to clarify this by means of a classical example. There are two possible interpretations of the sentence "Socrates is a man":

1. Socrates belongs to the class of all human beings;
2. Socrates exhibits the property of *being a man*;

Usually, in mathematics, the two views are assumed to be equivalent, and a predicate is taken as coinciding with the set of entities that satisfy it. This view is however too simplistic, since in Tarskian semantics set membership is taken as a basis to decide the *truth value* of property instantiation, so the former notion is independent from the latter. The existence of a mapping between the two relations does not justify their identification: one thing is a set, another thing is a property common to the elements of a set. A set may have many common properties, or maybe none[3]. A set has a cardinality, while a property abstracts from cardinality. A set is not something that can be multiply "present" in different things like a property: a set is a *particular*, a property is a *universal*. Membership involves the former, instantiation the latter.

So properties (universals) *correspond* to sets (called their *extension*), but are not sets. We may wonder however whether two universals that correspond to the same set are the same. Those who take intensionality into account usually refuse this assumption. The classical "realist example" of intensionality is that of the three predicates "human", "featherless biped", and "animal that laughs" that have the same extension but are considered to be different. An interesting alternative, suggested in [Lewis, 1983], is to include in the extension of a predicate all its possible instances (*possibilia*). In this case, "featherless biped" would include other instances besides humans, so that we can more safely assume that two universals are the same if they have the same extension.

A final problem concerns the possibility of having a (first order) logical theory of universals. In general, this appears to be impossible, since the predicates used to talk about universals (like instantiation) would themselves refer to universals. The solution we adopt is to reserve the term "universal" to those properties and relations whose instances are particulars. Limiting our domain to the first two levels, we can aim at building a separate meta-theory that accounts for the distinctions we need for our purposes. To stick to first order logic, however, we need to avoid quantifying on arbitrary universals. To this purpose, we adopt a practical suggestion proposed by Pat Hayes[4], to further restrict the universals we quantify on to a pre-defined set of relevant properties and relations, corresponding to the predicates explicitly mentioned in our object-level theory. Within this theory, we can state some minimal ground axioms for the instantiation relation, and introduce definitions for particulars and universals. Reading $I(x, y)$ as "x is an instance of y", we have:

(I1) $I(x, y) \rightarrow \neg I(y, x)$ (*asymmetry*)
(I2) $(I(x,y) \quad I(x,z)) \rightarrow (\neg I(y,z) \quad \neg I(z,y))$ (*antitransitivity*)

$Par(x) =_{\mathrm{def}} \neg \quad y(I(y, x))$
$Uni(x) =_{\mathrm{def}} \neg Par(x)$

We shall not discuss distinctions among universals in detail here. A preliminary discussion on this topic (focused on properties) has been published in [Guarino and Welty, 2000a].

4.2 Parthood

The *parthood* relation is a very basic and investigated notion, which has been formalized only at the beginning of 20th century [Leonard and Goodman, 1940; Lesniewski, 1991]. These works intend to build a single theory (called *classical extensional mereology*) that, unlike set theory, is founded only on concrete entities. More recently, [Simons, 1987] and [Casati and Varzi, 1999] pointed out that we can have different mereologies corresponding to different parthood relations, and made explicit the formal dependencies among them.

[2] The term "universal" is due to the fact that, metaphorically, we may see a property as multiply *present* in different things. Note however that this doesn't mean that different instances of the same universal have any *part* in common.

[3] In other words, a set doesn't coincide with its characteristic function.

[4] Message to the IEEE SUO list, http://suo.ieee.org

We shall write $P(x, y)$ as "*x is a part of y*". Only three ground axioms (P1-P3) are considered as minimal, although the *weak supplementation* axiom (P4) is often accepted:

(P1) $P(x, x)$
(P2) $(P(x, y) \quad P(y, x)) \rightarrow x = y)$
(P3) $(P(x, y) \quad P(y, z)) \rightarrow P(x, z)$
(P4) $PP(x, y) \rightarrow \quad z(P(z, y) \quad \neg O(z, x))$

where

(DPP) $PP(x, y) =_{\text{def}} (P(x, y) \quad \neg P(y, x))$
(DO) $O(x, y) =_{\text{def}} \quad z(P(z, x) \quad P(z, y))$.

The *extensionality* axiom (P5) and the stronger *supplementation* axiom (P6)[5]:

(P5) $(\quad z(PP(z, x)) \quad z(PP(z, x) \rightarrow PP(z, y))) \rightarrow P(x, y)$
(P6) $\neg P(x, y) \rightarrow \quad z(P(z, x) \quad \neg O(z, y))$

are much more controversial. It is safer therefore to assume they hold only for some classes of entities called *extensional*[6] entities.

Axioms guaranteeing existence of sum, difference, product, fusion of entities or establishing mereological properties (such as atomicity or divisibility) are debatable, and then we shall not commit to them at this stage of our methodology.

4.3 Connection

Parthood only is not enough to analyze the *internal structure* of a given entity, as it only allows us to check whether it is atomic or divisible. To the purpose of capturing at least some basic intuitions related to the notion of whole, connection is usually introduced in the meretopological literature [Simons, 1987; Varzi, 1999] as a further primitive in addition to parthood. It is assumed to satisfy the following minimal axioms:

Ground axioms:
(C1) $C(x, x)$
(C2) $C(x, y) \rightarrow C(y, x)$

Link with part relation:
(C3) $P(x, y) \rightarrow \quad z(C(z, x) \rightarrow C(z, y))$

Note that from (C3) and (P1) we can deduce (C2), then (C2) is redundant. The converse of (C3) is controversial, as it seems to be acceptable only for spatial regions.

These axioms can be specialized in various ways to account for different notions of connection. For instance, within topological connection between 3-d regions, it may be useful to distinguish among point-connection, line-connection, and surface-connection [Borgo *et al.*, 1996].

4.4 Location and Extension

Recently, Casati and Varzi [Casati and Varzi, 1999] have axiomatized the notion of location by means of a primitive *L* intended to capture the intuition of "being (exactly) in a place". Although their approach is focused on space only, we believe it can be generalized to account for the relationship existing between arbitrary entities and four-dimensional regions. Since – at least at this point – we want to be neutral about the commitments on the distinction between continuants and occurrents, we prefer renaming this relation in terms of *being extended in a (n-dimensional) region.* We introduce therefore a binary primitive $E(x, y)$ to be read as "*x* is the extension of *y*"[7], and we assume for it the axioms from [Casati and Varzi, 1999]:

Ground Axioms:
(E1) $(E(x, y) \quad E(z,y)) \rightarrow x = z$ (*functionality*)
(E2) $E(x, y) \rightarrow E(x, x)$ (*conditional reflexivity*)

Links with parthood:
(E3) $(P(x, y) \quad E(z, x) \quad E(w, y)) \rightarrow P(z, w)$
(E4) $(P(x, y) \quad E(y, z)) \rightarrow PE(x, z)$

where

$PE(x, y) =_{\text{def}} \quad z(P(z, y) \quad E(x, z))$ (*partial extension*)

Link with connection:

(E5) $(C(x, y) \quad E(z, x) \quad E(w, y)) \rightarrow C(z, w)$

Note that transitivity and antisymmetry follow from ground axioms. Note that we do not exclude that different entities can have the same extension, and that we assume a region as something that is extended in itself. Some useful definitions follow:

$Reg(x) =_{\text{def}} E(x, x)$ (*x is a region*)
$Extd(x) =_{\text{def}} \quad y(E(y, x))$ (*x is extended*)
$Coext(x, y) =_{\text{def}} \quad z,u(E(z, x) \quad E(u, y) \quad u = z)$ (*x and y are co-extensional*)
$OC_{\varphi}(x, y) =_{\text{def}} \varphi(x) \quad E(y, x) \quad z((\varphi(z) \quad E(y, z)) \rightarrow O(z, x))$ (*x φ-occupies y*)

We take for granted the further axioms introduced by [Casati and Varzi, 1999] to ensure the topological properties of regions (pp. 122-126), which will not be discussed here. On the basis of this theory, the following relevant theorem can be proved for any extensional property φ (see section 5.2):

$(OC_{\varphi}(x, y) \quad OC_{\varphi}(z, y)) \rightarrow z = x.$

We can't prove the same for co-extensionality, and this makes clear the difference between being extended in a region and occupying that region.

4.5 Dependence

We consider here ontological dependence as a general relation potentially involving all the entities of the domain. In this sense we try to understand dependence as a formal relation with a minimal ontological commitment, which can be specialized in different ways. The classical philosophical reference for the notion of dependence is Husserl's work

[5] Note that P6 implies P5, but not viceversa.
[6] Unfortunately, this adjective is used with different meanings in the literature, see section 5.2.
[7] Note that we reversed the arguments of Casati and Varzi's *L* primitive.

[Husserl, 1970]. Recently, Fine and Simons [Simons, 1987; Fine, 1995b] have suggested some alternative formalizations of Husserl's analysis, discussing their problems and possible solutions.

Following Husserl, Fine proposes four axioms for dependence:

Ground Axioms:
(D1) $D(x, x)$
(D2) $D(x, y) \quad D(y, z) \rightarrow D(x, z)$

Links with Part relation:
(D3) $P(x, y) \rightarrow D(y, x)$
(D4) $\quad y(D(x, y) \quad z(D(x, z) \rightarrow P(z, y)))$

Is this a good axiomatization of dependence relation? A minor problem is that (D1) is provable from (D3) and from the reflexivity of parthood, (P1), then (D1) is redundant. Another problem regards (D4). This axiom guarantees the existence of an entity y that is the maximal (with respect to part relation) entity from which x depends, and then it is clearly not ontologically neutral.

Moreover, Simons criticizes these axioms from a more general point of view. He points out that these axioms can be interpreted in terms of weak topological structures, where dependent entities correspond to non-closed sets, and independent entities correspond to closed sets. Dependence would therefore resemble a sort of topological relation, and this may sound as counterintuitive.

Indeed, ontological dependence is usually not introduced as a primitive relation, but rather defined in terms of a modal operator and an existence predicate (Ex) as:

(DD) $\quad D(x, y) =_{\text{def}} \Box(Ex(x) \rightarrow Ex(y))$.

In order to accept this, we have to accept however that dependence is intrinsically linked to modality, and somebody finds this debatable, too. If we want to be neutral with respect to this issue, we need a theory that is compatible with the modal interpretation of D relation. But, as Simons points out, if we interpret D as in (DD), axiom (D3) is satisfied only if we either subscribe to mereological essentialism (any part of x is necessarily such) or if we consider a modal interpretation of P (part means *essential part*). Otherwise in general from "x is part of y it does not follow that y could not exist without x" (Simons, p. 317). This is a big problem. If we abandon axioms (D3- D4) the characterization of dependence relation is really weak.

We are tempted therefore to accept (DD). In this case we have however other problems. One problem is that we may have different kinds of modal operators each inducing different kinds of dependence relations and that present technical difficulties (for example, *formal necessity*, *material necessity*, *nomological necessity*, etc.). A more serious problem is the characterization of predicate Ex. This seems really not so simple. For example, does *being something* coincide with *existing*? Do things like ordinary objects and events exist in the same way? (see [Fine, 1995a]). In order to clarify these issues we may introduce time too, but in this case we need either to introduce another modal operator that interacts with the first one, or to treat time independently. The latter approach has been adopted in [Thomasson, 1999], where the author informally introduces different kinds of temporal dependence, such as:

$CD(x, y) =_{\text{def}} \Box(Ex(x, t) \rightarrow Ex(y, t))$ (*constant dep.*)
$HD(x, y) =_{\text{def}} \Box(Ex(x, t) \rightarrow (Ex(y, t') \quad t' \leq t)$ (*historical dep.*)

Thomasson also includes the possibility for a universal to depend on a specific particular (being a wife of Henry VIII depends on Henry VIII), and for a particular to depend only *generically* on another particular that instantiates a specific universal (the US generically depend on some US citizen).

We find these definitions extremely interesting intuitively, but we do not attempt at formalizing them here. So, for the time being, we take only axioms (D1) and (D2), leaving the interpretation of ontological dependence to intuition. We introduce however some useful definitions based on P and D:

$MD(x, y) =_{\text{def}} D(x, y) \quad D(y, x)$ (*mutual dependence*)
$SD(x, y) =_{\text{def}} D(x, y) \quad \neg D(y, x)$ (*one-side dependence*)
$ED(x, y) =_{\text{def}} D(x, y) \quad \neg P(y, x)$ (*external dependence*)

5 Formal Properties

On the basis of the formal relations discussed above, let us briefly introduce a set of formal properties that we believe especially useful for our purposes. For the sake of simplicity, our domain of quantification will be limited to particulars so that the formal properties will not correspond to logical definitions, but will be stated in the meta-language. Those meta-level definitions that classify a particular with respect to the universal denoted by φ are expressed by using a φ subscript.

5.1 Concreteness and abstractness

In section 4.4 we have already defined the notion of an extended entity as something that extends in a (spatiotemporal) region. We shall take the property of being extended as synonymous of *being concrete.* A non-extended entity will be called *abstract*.

Note that this sense of "abstract" has nothing to do with the process of abstracting a common property from a set of entities. So the decision whether properties (or universals) are abstract or concrete, according to our terminology, cannot be taken on the basis of the theory of extension we have introduced. What the theory tells us is that, if the elements of a set or the instances of a property are concrete, we assume universals to be concrete, then we have to interpret the meaning of parthood and connection for universals in a suitable way. Moreover, we have to establish a link between the extension of a particular and the extension of the universal that it instantiates. Similar difficulties would occur assuming that sets are concrete, since in this case we need a theory that links their extension to that of their members (and to the parts of their members). For these reasons, it seems pretty safe to stick to the usual assumption that universals and sets are both abstract. Collections, which will be discussed below, are the concrete correspondent of sets.

5.2 Extensionality

We say that an entity is *extensional* if and only if everything that has the same proper parts is identical to it:

$Extl(x) =_{def} \; z(PP(z, x) \quad (\; z(PP(z, x) \leftrightarrow PP(z, y)) \rightarrow x = y)$

Examples of extensional entities are regions and amounts of matter.

We say that a property is *extensional* iff all its instances are extensional. We say in this case that this property *carries an extensional criterion of identity*[8].

Unfortunately, the adjective "extensional" is also used with different meanings in the literature. Sets are said to be extensional since they are identical when they have the same members, and properties are considered as extensional when properties with the same instances are taken as identical.

5.3 Unity and plurality

We believe that the formal relations we have introduced allow us to exactly define the notion of unity, but this requires some care.

Let us first give some definitions based on the parthood relation, which may capture some notions related to that of unity:

$At(x) =_{def} \neg \; y(PP(y, x))$ (*atomicity*)
$Div(x) =_{def} \neg At(x)$ (*divisibility*)
$At_\varphi(x) =_{def} \varphi(x) \quad \neg \; y(\varphi(y) \quad PP(y, x))$ (*φ-atomicity*)
$Div_\varphi(x) =_{def} \varphi(x) \quad y(\varphi(y) \quad PP(y, x))$ (*φ-divisibility*)
$IHom_\varphi(x) =_{def} \varphi(x) \quad y(PP(y, x) \rightarrow \varphi(y))$ (*φ-int. homogeneity*)
$Max_\varphi(x) =_{def} \varphi(x) \quad \neg \; y(\varphi(y) \quad PP(x, y))$ (*φ maximality*)
$EHom_\varphi(x) =_{def} \varphi(x) \quad y(PP(x, y) \rightarrow \varphi(y))$(*φ-ext. homogeneity*)
$\Sigma_\varphi(x) =_{def} \; y(P(y, x) \rightarrow \; z(\varphi(z) \quad P(z, x) \quad O(z, y))$(*sum of φs*)

The notion of maximality seems indeed very much related to unity (wrt a certain φ), but it does not account for the way the various parts of x are bound together. Indeed, there are different aspects behind the notion of unity of an object, which are merged together in the following definition:

> "Every member of some division of the object stands in a certain relation to every other member, and no member bears this relation to anything other than members of the division." ([Simons, 1987], p.328)

We see here at least three fundamental aspects: a notion of division within a whole, with *members* of such division; a suitable *unifying relation* that binds the members together, and a *maximality constraint* with respect to this relation on the members. The notion of "member of a division" is the subtle issue here. Simons takes class membership as a primitive distinct from parthood. We'd rather analyze it in terms of parthood, in the spirit of the analysis presented in [Guarino and Welty, 2000b]. The definition of unity proposed in that paper has however some problems[9], so we propose here a new one that captures more carefully the notion of member of a division.

Our intuition is that any member of x is a *special part* of x. So we need a property that individuates the members within the parts of x. Observe that, if x forms a unity under a unifying relation R, then the property we need must only pick up the parts of x that belong to R's domain. All other parts can be ignored[10].

Note now that R must be at least symmetric and reflexive (let's postpone by now the discussion about transitivity). Then we can define a predicate δ_R denoting R's domain, which must hold when x is a member of a division unified by R:

$$\delta_R(x) =_{def} R(x, x)$$

We can observe that, if R is defined on the whole domain, then $\delta_R(y)$ also holds for all the parts y of x.

We define now the notion of a whole as follows:

$\upsilon_R(x) =_{def} \Sigma_{\delta_R}(x) \quad y,z((\delta_R(y) \quad \delta_R(z) \quad P(y, x) \quad P(z, x)) \rightarrow R(y, z))$ (*x is unified by R*)
$\omega_R(x) =_{def} Max_{\upsilon_R}(x)$ (*x is a whole under R*)

The first definition says that x is unified by R iff it is a sum of entities belonging to R's domain, and all these entities are linked together by R. The second one says that x is a whole under R iff it is *maximally* unified by R.

Let us discuss now the assumptions regarding R's transitivity. At a first sight, it would be obvious to assume R as transitive; together with the previous assumptions, this would result in R being an equivalence relation. However, this would exclude the possibility of *overlapping wholes* with a *common* unifying relation[11]. Consider for example the notions of *committee* or *organization*: two committees may have a member in common while being two different wholes. Of course, in a strict sense, there would two different unifying relations in this case (say, having mission A vs. having mission B). The point is that there would be no *common* unifying relation attached to the property *committee*. A plausible common relation would be "having the same mission", but this is not transitive. This is why, con-

[8] An extensive analysis of criteria of identity has been done elsewhere [Guarino and Welty, 2000b; Guarino and Welty, 2001]; for our purposes, we shall only distinguish here between extensional and non-extensional identity criteria.

[9] Consider the following counterexample: suppose you want to say that all the children *a, b, c* of a certain person form a whole. So all the parts of $a+b+c$ must be linked together by the unifying relation "having the same parent". But two of them, namely a+b and b+c, are not linked by such relation, since they are not persons. Another problem is linked to the fact that the previous definition excludes the possibility of overlapping of entities that are wholes (see below).

[10] We may also study the property of *internal uniformity* of x with respect the predicate $\varphi(y) =_{def} R(y, x)$ but this is another problem.

[11] We are grateful to Aaron Kaplan for this counterexample.

trary to the previous papers by Guarino and Welty, we shall not assume transitivity for *R*.

For the purpose of ontological analysis, it is interesting to explore how various unifying relations can be defined on the basis of simpler relations, called *characteristic relations* ([Simons, 1987] p.330). This analysis can be used to introduce different kinds of wholes.

In particular, from the cognitive point of view, it is very interesting to consider topological connection as a characteristic relation. More exactly, the cognitively relevant characteristic relations are those that restrict topological connection to hold between physical entities of the same kind, such as matter, color, or physical bodies (otherwise, using topological connection only, the only whole would be the universe).

Under this perspective, take C_ϕ as the transitive closure of the projection of C on ϕ entities. We say that x is a *topological whole under* ϕ if $\omega_{C_\phi}(x)$.

We can now introduce the notions of *singularity* and *plurality*, assuming that they are cognitively bound to topological connection:

$$Sing_\phi(x) =_{\mathrm{def}} \omega_{C_\phi}(x) \qquad (\textit{singularity})$$
$$Plur_\phi(x) =_{\mathrm{def}} \neg Sing_\phi(x) \quad y(PP(y, x) \quad Sing_\phi(y)) \qquad (\textit{plurality})$$

A singular entity is therefore one that is a topological whole. We define a plurality as anything that contains a topological whole and is not itself a topological whole.

Topological wholes have two parameters, corresponding to the C and ϕ above. If we take C as the usual topological connection, an isolated piece of matter will be a topological whole under "matter", while a spot of color will be a topological whole under "color". Note that nothing excludes a topological whole (under a certain kind of connection) to include other topological wholes (under a different kind of connection): think of a lump of spheres, which can be seen as a whole under point connection and contains many wholes under surface connection.

Note that if something does not contain a whole, it will be neither singular nor plural (think for instance of an undetached piece of matter).

A special case of plurality is a *collection*, which must be a sum of wholes. Each of these wholes will be a member of the collection.

Within singular entities, it may be interesting to distinguish between homogenous and non-homogeneous entities with respect to φ:

$$Simple_{\phi\varphi}(x) =_{\mathrm{def}} Sing_\phi(x) \quad IHom_\varphi(x)$$
$$Complex_{\phi\varphi}(x) =_{\mathrm{def}} Sing_\phi(x) \quad \neg IHom_\varphi(x)$$

For instance, if we assume singularity as based on point connection (that is, roughly, physical contact), we have that a physical body is homogenous wrt surface-self-connection, while an assembly formed by different bodies that touch each others is not.

5.4 Dependence and Independence

Finally, the last formal property that we consider is whether or not an entity is *externally* dependent, i.e. dependent on other things besides its parts:

$$Dep(x) =_{\mathrm{def}} \quad y(ED(x, y)) \qquad (\textit{dependence})$$
$$Dep_\varphi(x) =_{\mathrm{def}} \quad y(\varphi(y) \quad ED(x, y)) \qquad (\varphi\text{-}\textit{dep., or generic dep.})$$
$$Ind(x) =_{\mathrm{def}} \neg Dep(x) \qquad (\textit{independence})$$

According to our discussion in section 4.5, we have however to select an intended interpretation for *D*, since there are different kinds of dependence. For the purpose of isolating broad, relevant categories of entities, we believe that a special importance should be given to what Thomasson calls *constant dependence*, whose proper formalization requires an account of time that must be subject of future work. Under this view, for example, we can stipulate that ordinary objects (*continuants*) are independent, while events (*occurrents*) are dependent. More work on this is needed, however.

Conclusions

Developing a well-founded top-level ontology is an very difficult task, that requires a carefully designed methodology and rigorous formal framework. We hope to have contributed on both these aspects.

Since this is work in progress, we haven't been able to explore and discuss in detail the practical consequences of the methodology we have presented, although we have definite evidence of its relevance.

We are presently at step 4 of the sequence discussed in section 3.2. We hope in the possibility of a cooperative effort to proceed (through refinements and adjustments) in the way we have outlined.

This work has been done in the framework of the European Eureka project E!2235 "IKF" (Intelligent Knowledge Fusion). In this framework, we plan to develop a general reference ontology linked to a lexical resource such as WordNet, by using the methodology we have outlined. The final result will be of public domain, and will hopefully profit from (and contribute to) existing cooperation initiatives in this area, such as the IEEE SUO.

Bibliography

[Borgo *et al.*, 1996] Stefano Borgo, Nicola Guarino and Claudio Masolo. A Pointless Theory of Space based on Strong Connection and Congruence. In *Principles of Knowledge Representation and Reasoning (KR96)*, pp. 220-229, Boston, MA, 1996.

[Casati and Varzi, 1999] Roberto Casati and Achille Varzi. *Parts and Places.* MIT Press, Cambridge, 1999.

[Fellbaum, 1998] C. Fellbaum (ed.), *WordNet. An Electronic Lexical Database*. MIT Press, Cambridge, 1998.

[Fine, 1995a] K. Fine. Ontological Dependence. *Proceedings of the Aristotelian Society*, 95:269-90,1995a.

[Fine, 1995b] K. Fine. Part-Whole. In B. Smith and D.W. Smith (eds.), *The Cambridge Companion to Husserl*, pp. 463-485, New York, Cambridge University Press, 1995b.

[Gibson, 1977] J. J. Gibson. The Theory of Affordances. In R. E. Shaw, J. Bransford (eds.), *Perceiving, Acting and Knowing*, Hillsdale, LEA, 1977.

[Guarino and Welty, 2000a] Nicola Guarino and Christopher Welty. A Formal Ontology of Properties. In *Proceedings of the ECAI-00 Workshop on Applications of Ontologies and Problem Solving Methods*, pp. 12-1 12-8, Berlin, Germany, 2000a.

[Guarino and Welty, 2000b] Nicola Guarino and Christopher Welty. Identity, Unity, and Individuality: Towards a Formal Toolkit for Ontological Analysis. In Werner Horn (ed.) *ECAI-2000: Proceedings of the 14th European Conference on Artificial Intelligence*, pp. 219-223, IOS Press, Berlin, Germany, 2000b.

[Guarino and Welty, 2001] Nicola Guarino and Christopher Welty. Subsumption and Identity. Technical Report n. 01/2001, LADSEB-CNR, Italy, 2001.

[Husserl, 1970] Edmund Husserl. *Logical Investigations.* Routledge and Kegan Paul, London, 1970.

[Kaplan, 2001] A. N. Kaplan, *Personal Communication*. 2001.

[Lakoff and Johnson, 1999] G. Lakoff and M. Johnson. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books, 1999.

[Leonard and Goodman, 1940] H. Leonard, S. and N. Goodman. The Calculus of Individuals and Its Uses. *Journal of Symbolic Logic*, 5:45-55,1940.

[Lesniewski, 1991] S. Lesniewski. *Collected Works.* Kluwer, Dordrecht, 1991.

[Lewis, 1983] David Lewis. New Work for a Theory of Universals. *Australasian Journal of Philosophy*, 61(4)1983.

[Simons, 1987] P. Simons. *Parts: a Study in Ontology.* Clarendon Press, Oxford, 1987.

[Smith, 1998] Barry Smith. Basic Concepts of Formal Ontology. In Nicola Guarino (ed.) *Formal Ontology in Information Systems*, pp. 19-28, IOS Press, Amsterdam, 1998.

[Thomasson, 1999] Amie L. Thomasson. *Fiction and Metaphysics.* Cambridge University Press, Cambridge, 1999.