# A Formal Theory of Conceptual Modeling Universals

Giancarlo Guizzardi[1], Gerd Wagner[2], Marten van Sinderen[1]

[1] Centre for Telematics and Information Technology
University of Twente
Enschede, The Netherlands
{guizzard@cs.utwente.nl, sinderen@ctit.utwente.nl}

[2] Eindhoven University of Technology,
Faculty of Technology Management,
Eindhoven, The Netherlands
G.Wagner@tm.tue.nl
http://tmitwww.tm.tue.nl/staff/gwagner

**Abstract**. Conceptual Modeling is a discipline of great relevance to several areas in Computer Science. In a series of papers [1,2,3] we have been using the General Ontological Language (GOL) and its underlying upper level ontology, proposed in [4,5], to evaluate the ontological correctness of conceptual models and to develop guidelines for how the constructs of a modeling language (UML) should be used in conceptual modeling. In this paper, we focus on the modeling metaconcepts of *classifiers* and *objects* from an ontological point of view. We use a philosophically and psychologically well-founded theory of universals to propose a UML profile for Ontology Representation and Conceptual Modeling. The formal semantics of the proposed modeling elements is presented in a language of modal logics with quantification restricted to Sortal universals.

## 1 Introduction

Conceptual Modeling is regarded as a discipline whose importance spreads throughout several areas in the realm of Computer Science (e.g. Software Engineering, Information Systems Design, Domain Engineering, Database Design, Requirements Engineering, and Knowledge Engineering, among others). Its main objective is concerned with identifying, analyzing and describing the essential concepts and constraints of a universe of discourse with the help of a (diagrammatic) modeling language that is based on a small set of basic meta-concepts (forming a metamodel). Ontological modeling, on the other hand, is concerned with capturing the relevant entities of a domain in an ontology of that domain using an ontology specification language that is based on a small set of basic, domain-independent ontological categories (forming an upper level ontology). While conceptual modeling languages are evaluated on the basis of their successful use in the practice domain information modeling, ontology specification languages and their underlying upper level ontologies have to be rooted in principled philosophical theories about what kinds of things exist and what their basic relationships with each other are.

The Unified Modeling Language (UML) is a language initially proposed as a unification of several different visual notations and modeling techniques used for systems design [6]. UML is now a *de facto* standard for modeling computational systems and, recently, it has been proposed that the language should be also used as an Ontology Representation Language [7]. Moreover, in this paper the authors argue that although UML lacks a precise definition of its formal semantics, this difficulty shall be overcome with the current developments made by the precise UML community[1].

We believe, however, that defining constructs of a conceptual modeling only in terms of its mathematical semantics, although essential, it is not sufficient to make it a suitable ontology representation language. The position defended here is that, in order to model reality, a conceptual modeling language should be founded on formal upper-level ontologies. In other words, it should have both, formal and ontological semantics.

In a series of papers we have been employing the General Ontological Language (GOL) and its underlying upper level ontology, proposed in [4,5], to evaluate the ontological correctness of UML conceptual models and to develop guidelines that assign well-defined ontological semantics to UML

[1] http://www.cs.york.ac.uk/puml/

modeling constructs. In [1], we have discussed the meaning of the UML metaconcepts of *classes and objects*, *powertypes*, *association* and *part-whole* relations (*aggregation/composition*). The UML metaconcepts of *abstract classes* and *datatypes* are addressed in a companion paper [2]. The work presented here can be seen as a continuation of this work in which we focus on one aspect of the philosophical problem between universals and particulars (roughly, classifiers and object instances in UML terms).

Although classifier modeling constructs are fundamental in conceptual modeling (being present in all major conceptual modeling languages) there is still a deficiency of methodological support for helping the user of the language deciding how to model the elements of a given domain. In practice, a set of primitives is often used to model distinctions in different types of classifiers (Type, Role, State, Mixin, among others). However, the choice of how the elements that denote universal properties in a domain (viz. Person, Student, Red Thing, Physical Thing, Deceased Person, Customer) should be modeled is often made in ad hoc manner. Likewise, it is the judgment of what are the admissible relations among these modeling elements.

This paper proposes a philosophically and psychologically well-founded theory of universals (section 2). This theory is further used to generate a typology of UML classifiers together with a set of methodological guidelines that governs its use (section 3). Additionally, we provide a formal characterization of the types of universals proposed in section 2 in a language of modal logics with restricted quantification (section 4). Finally, section 6 elaborates on some conclusions and future work.

## 2 Towards a theory of classifier types for Conceptual Modeling: philosophical and psychological foundations

In [8], van Leeuwen shows an important syntactical difference in natural languages that reflects a semantical and ontological one, namely, the difference between common nouns (CNs) on one side and arbitrary general terms (adjectives, verbs, mass nouns, etc…) on the other. CNs have the singular feature that they can combine with determiners and serve as argument for predication in sentences such as: (i) *(exactly) five mice were in the kitchen last night; (ii) the mouse which has eaten the cheese, has been in eaten in turn by the cat.*

In other words, if we have the patterns *(exactly) five X…* and *the Y which is Z…*, only the substitution of X,Y,Z by CNs will produce sentences which are grammatical. To see that, we can try the substitution by the adjective *Red* in the sentence (*i*)*: (exactly) five red were in the kitchen last night.* A request to 'count the red in this room' cannot receive a definite answer: Should a red shirt be counted as one or should the shirt, the two sleeves, and two pockets be counted separately so that we have five reds? The problem in this case is not that one would not know how to finish the counting but that one would not know how to start since arbitrarily many subparts of a red thing are still red.

The explanation for this feature unique of CNs lies on the function that determinates (demonstratives and quantifiers) play in noun phrases, which is to determine a certain range on individuals. Both reference and quantification requires that the thing (or things) which are referred or which form the domain of quantification are determinate individuals, i.e., their conditions for *individuation* and *identity* must be determinate. In other words, if it is not determinate how to count Xs or how to identify X that is the same as Y, the sentences in the patterns (i) and (ii) do not express determinate propositions, i.e. propositions with definite truth values.

The distinction between the grammatical categories of CNs and arbitrary general terms can be explained in terms of the ontological categories of Sortal and Characterizing universals [9], which are roughly their ontological counterparts. Whilst the latter supply only a principle of application for the individuals they collect, the former supply both a principle of application and a principle of identity. A principle of application is that in accordance with which we judge whether a general term applies to a particular (e.g. whether something is a Person, a Dog, a Chair or a Student). A principle of identity supports the judgment whether two particulars are the same, i.e., in which circumstances the identity relation holds.

In [10], Macnamara, investigates the role of sortal concepts in cognition and provides a comprehensive theory for explaining the process that a child undergoes when learning proper nouns and common nouns. He proposes the following example: suppose a little boy (Tom), which is about to learn the meaning of a proper name for his puppy. When presented to the word Spot", Tom has to decide what it refers to. One should notice that a demonstrative such as "that" will not be sufficient to determinate the bearer of the proper name? How to decide that "that" which changes all its perceptual properties is still Spot? In other words, which changes can Spot suffer and still be the same? As Macnamara (among others) shows, answers to these questions are only possible if Spot is taken to be a proper name for an individual, which is an instance of a Sortal universal. The principles of identity supplied by the Sortals are essential to judge the validity of all identity statements. For example, if for an instance of the sortal *Statue* loosing a piece will not alter the identity of the object, the same does not hold for an instance of *Lump of Clay*.

The statement that we can only make identity and quantification statements in relation to a Sortal amounts to one of the best-supported theories in the philosophy of language, namely, that the identity of an individual can only be traced in connection with a Sortal Universal, which provides a *principle of individuation* and *identity* to the particulars it collects [8,10,11,12]. The position advocated in this article affirms an equivalent stance for a theory of conceptual modeling. We defend that among the conceptual modeling counterparts of general terms (classifiers), only constructs that represent substance sortals can provide a principle of identity and individuation for its instances. As a consequence, the following principle can be postulated:

**<u>Postulate 1:</u>** Every Object in a conceptual model (CM) of the domain must be an instance of a CM-class representing a sortal.

As argued by Kripke [13], a proper name is a rigid designator, i.e. it refers to the same individual in all possible situations, factual or counterfactual. For instance, it refers to the individual Mick Jagger both now (when he is the lead singer of Rolling Stones and 60 years old) and in the past (when he was the boy Mike Philip living in Kent, England). Moreover, it refers to the same individual in counterfactual situations such as the one in which he decided to continue in the London School of Economics and has never pursued a musical career. We would like to say that the boy Mike Philip is identical with the man Mick Jagger that he latter became. However, as pointed out by Wiggins [14] and Perry [15], statements of identity only make sense if both referents are of the same type. Thus, we could not say that a certain Boy is the same Boy as a certain Man since the latter is not a Boy (and vice-versa). However, as Putnam put it, when a man x points to a boy in a picture and says "I am that boy", the pronoun "I" in question is typed not by Man but by a supertype of Man and Boy (namely, Person) which embraces x's entire existence [16]. A generalization of this idea amount to a thesis, proposed by Wiggins, named thesis D [14]: If an individual falls under two sortals in the course of its history there must be exactly one ultimate sortal of which both sortals are specializations. Griffin elaborates Wiggins' thesis D in terms of two correlated principles:

**a)** **<u>The Restriction Principle:</u>** if an individual falls under two distinct sortals F, F' in the course of its history then there is at least one sortal which F and F' are both specializations.

**b)** **<u>The Uniqueness Principle</u>:** if an individual falls under two distinct sortals F, F' in the course of its history then there is only *one ultimate sortal* which F and F' are both specializations. A sortal F is ultimate if there is no other sortal F' distinct from F which F specializes.

It is not the case that two incompatible principles of identity could apply to the same individual x, otherwise x would not be a viable entity (determinate particular) [8]. Imagine an individual x which is an instance of both Statue and Lump of clay. Now, the answer to the question whether loosing a piece will alter the identity of x is indeterminate since each of the two principles of identity that x obeys imply a different answer. As a consequence, we can say that if two sortals F and F' intersect (i.e. have common individuals in their extension), the principles of identity contained in them must be equivalent. Moreover, F and F' cannot supply a principle of identity for x, since both sortals apply to x only contingently and a principle of identity must be used to identify x all possible worlds. Therefore, there must be a sortal G that supplies the principle of identity carried by F and F'. This proves the restriction principle. The uniqueness of the ultimate sortal G can be argued as follows: (i) G is a sortal, since it supplies a principle of identity for all the things in its extension; (ii) if it restricts a sortal H then, since H cannot supply a incompatible principle of identity, H either: is equivalent to G (i.e. supply the same principle of identity) and therefore should be ultimate or does not supply a principle of identity for the particulars in its extension (see text on dispersive classifiers below). This proves the uniqueness principle. The unique ultimate sortal G that supplies the principle of identity for its instances is named a *substance sortal*.

As a consequence of the *uniqueness principle* we define a second postulate:

**<u>Postulate 2:</u>** An Object in a conceptual model of the domain cannot instantiate more than one CM-Class representing an ultimate Substance Sortal.

In the example above, the sortal Person is the *unique substance sortal* that defines the validity of the claim that Mick Jagger is the same as Mike Philip or, in other words, that Mike Philip persists through changes in height, weight, age, residence, etc… as the same individual. Person can only be the sortal that supports the proper name Mick Jagger in all possible situations because it applies necessarily to the individual referred by the proper name, i.e. instances of Person cannot cease to be so without ceasing to exist. As a consequence, the extension of a substance sortal is world invariant. This meta-property of classifiers is named *Modal Constancy* [12] or *rigidity* [17].

Sortals such as Boy and Adult Man in the example above, but also Student, Employee, Caterpillar and Butterfly, Philosopher, Writer, Alive and Deceased, which possibly apply to a continuant during a certain phase of its existence, are named phased-sortal in [14]. As a consequence of the *Restriction Principle* we have that for every phased-sortal PS that applies to a continuant, there is a substance sortal S of which PS is a specialization.

Contrary to substance sortals, phased-sortals apply to individuals contingently and, thus, do not enjoy modal constancy. For example, for an individual John instance of *Student*, we can easily imagine John moving in an out of the *Student* type, while being the same individual, i.e. without loosing his identity. Moreover, for every instance x of Student in a world w, there is another world w' in which x is not an instance of Student. This meta-property of classifiers is name *anti-rigid* in [17].

By considering how these different universals stand w.r.t rigidity another postulate can be derived:

**Postulate 3:** A CM-Class representing a rigid classifier cannot be a subclass a CM-Class representing an anti-rigid classifier

If PS is a phased-sortal and S is the substance sortal specialized by PS, there is a specialization condition $\phi$ such that x is a PS iff x is a S that satisfies $\phi$ [10]. A further clarification on the different types of specialization conditions allows us to distinguish between two different types of phased-sortals which are of great importance to the practice of conceptual modeling, namely, *phases* and *roles*.

Phases constitute possible stages in the history of a substance sortal. Examples are: (a) Alive and Deceased: as possible stages of a Person; (b) Catterpillar and Butterfly of a Lepidopteran; (c) Town and Metropolis of a City; (d) Boy, Male Teenager and Adult Male of a Male Person. *Classifiers representing phases constitute a partition of the substance sortal they specialize.* For example, if <Alive,Deceased> is a *phase-partition* of a sustance sortal Person then for every world w, every Person x is either and instance of Alive or of Deceased but not of both. Moreover, if x is an instance of Alive in world *w* then there is world *w*' such that x is not an instance of Alive in *w*', which in this case, implies that x is an instance of Deceased in *w*'.

Contrary to phases, roles do not necessarily form a partition of substance sortals. Moreover, they differ from phases in terms of the specialization condition $\phi$. For a phase P, $\phi$ represents a condition that depends solely on intrinsic properties of P. For instance, one might say that if Mick Jagger is a Living Person then he is a Person who has the property of being alive or, if Spot is a Puppy then it is a Dog which has the property of being less than a year old. For a role R, conversely, $\phi$ depends on extrinsic (relational) properties of R. For example, one might say that John is a Student then John is a Person who is enrolled in some educational institution or that, if Peter is a Customer then Peter is a Person who buys a Product y from a Supplier z. In other words, an entity plays a role in a certain context, demarcated by its relation with other entities.

Although Frege argued at length that "one cannot count without knowing what to count", in artificial logical languages inspired by him, natural language general terms such as CNs, adjectives and verbs are treated uniformly as predicates. For instance, if we want to represent the sentence "there are tall men", in the fregean approach of classical logic we would write $\exists$x Man(x) $\wedge$ Tall(x). This reading puts the count noun Man (which denotes a Sortal) on an equal logical footing with the predicate Tall. Moreover, in this formula, the variable x is interpreted into a "supposedly" universal kind Thing. So, the natural language reading of the formula should be "there are things which have the property of being a man and the property of being tall". Since, by postulate 1, all individuals must be instances of a substance sortal we must conclude that Thing is a unique universal ultimate sortal which is able to supply a principle of identity for all elements that we consider in our universe of discourse. Moreover, by postulate 2, this principle of identity must be unique. Can that be the case?

In [20], Hirsch argues that concepts such as Thing, (Entity, Element, among others) are *dispersive*, i.e. they cover many concepts with different principles of identity. For instance, in the extension of Thing we might encounter an individual x which is a cow and an individual y which is a watch. Since the principles of identity for Cows and Watches are not the same we conclude that Thing cannot supply a principle of identity for its instances. Otherwise, x and y would obey incompatible principles of identity and, thus, would not be determinate individuals. Therefore, as defended in [8,11,12,18], dispersive concepts do not denote sortals (despite the fact that they are considered CNs in natural languages) and *therefore cannot have direct instances.* More than that, since a principle of identity supplied by a substance sortal G is inherited by all classifiers that specialize G or, to put in another way, all subtypes of G carry the principle of identity supplied by G. Thus, all subclasses of a sortal are themselves sortals, ergo,

**Postulate 4:** A CM-Class representing a dispersive universal cannot be a subclass of a CM-Class representing a Sortal

# 3 An Ontologically well-founded UML profile for conceptual modeling

The Unified Modeling Language (UML) has built in extension mechanisms that allow one to modify the language elements to suite certain modeling needs. A coherent set of such extensions, defined accordingly to a specific purpose or domain, constitutes a *UML profile* [6].

A Stereotype is a lightweight extension mechanism that allows one to specialize UML modeling elements by defining additional constraints and sometimes a different graphical notation, so that they behave in some aspects as if they were instances of elements defined in new virtual metamodel. Stereotypes are also used to indicate difference in meaning or usage between modeling elements with a similar structure.

In [3], we have proposed a profile for UML to support the design of ontologically well-founded conceptual models according to the theory proposed in section 2. This profile (summarized in the table below) comprises of a set of stereotyped classes (specializations of the meta-construct class) that represents finer-grained distinctions between different types of substantial universals. Additionally, the profile incorporates a number of constraints that is applied to relations involving these stereotyped classes.

| Stereotype | Description | Constraints |
|---|---|---|
| «kind» A | A kind represents a *substance sortal*, i.e. rigid, externally independent universals that supply a principle of identity for its instances. Examples could be instances of Natural Kinds (such as Person, Dog, Tree) and artifacts (Chair, Car, Television). | Every object in conceptual model using this profile must be an instance of a Kind, directly or indirectly (postulate 1). Moreover, it cannot be an instance of more than one ultimate Kind (postulate 2). A supertype of a kind cannot be a member of {« subkind », « phase », « role », « roleMixin »} |
| «subkind» A<br>A | A subkind is a rigid, externally independent restriction of a kind which carries the principle of identity supplied by the kind. An example could be the subkind MalePerson of the kind Person. In general, the stereotype «subkind» can be omitted in conceptual models without loss of clarity. | A sypertype of a subkind cannot be a member of {« phase », « role », « roleMixin »} |
| «phase» A | It represents the phased-sortals *phase,* i.e. *anti-rigid* and *externally independent* universals defined as part of a partition of a kind. For instance, the partition {Catterpillar, Butterfly} of the kind Lepdopterum. | The phases $\{P_1 \ldots P_n\}$ that form a partition of a Kind K are defined in UML as a disjoint and complete generalization set. The kind K is always depicted as an abstract class. |
| «role» A | It represents a phased-sortal *role*, i.e. *anti-rigid* and *externally dependent* universal. For instance, the role student played by instance of the kind Person. | Roles and Phases are anti-rigid universals and cannot appear in a conceptual model as a superclass of a Kind (postulate 3). Moreover: Let X be a class stereotyped as « role » and r be an association representing X's restriction condition. Then, $\#X.r \geq 1$ |
| «category» *A* | It represents a rigid and externally independent ***non-sortal***, a dispersive universal that aggregates essential properties which are common to different kinds. For example, the category RationalEntity as a generalization of Person and IntelligentAgent. | A category cannot have direct instances and must be depicted as an abstract class. A supertype of a category cannot be a member of {« kind », « subkind », « phase », « role », « roleMixin »} |
| «roleMixin» *A* | It represents an anti-rigid and externally dependent ***non-sortal***, a dispersive universal that aggregates properties which are common to different roles. It includes formal roles such as *whole*/*part* and *initiatior/ responder*. | A role mixin cannot have direct instances and must be depicted as an abstract class. A supertype of a role mixin cannot be a member of {« kind », « subkind », « phase », « role »}. Let X be a class stereotyped as « roleMixin » and r be an association representing X's restriction condition. Then, $\#X.r \geq 1$ |

| | | |
|---|---|---|
| «mixin» **A** | The stereotype «mixin» represents properties which are essential to some of its instances and accidental to others (a meta-property named semi-rigidity in [17]). An example is the mixin *Seatable*, which represents a property that can be considered essential to the kinds Chair and Stool but accidental to Crate, Paper Box or Rock. | A mixin cannot have direct instances and must be depicted as an abstract class. A supertype of a mixin cannot be a member of {« kind », « subkind », « phase », « role », « roleMixin »} |

# 4 A formal characterization of the proposed categories

In this section we provide a formal characterization of the notions discussed in section 2 by using a language L of quantified modal logics with identity. A model M in this language is a structure $<W,D,\delta>$ where W is a non-empty set of worlds, D is a non-empty domain of objects and $\delta$ is an interpretation function assigning values to the non-logical constants of the language. The domain D of quantification is that of *possibilia*, which includes all possible entities independent of their actual existence. Therefore we shall quantify over a constant domain in all possible worlds. Moreover, all worlds are equally accessible and therefore we omit the accessibility relation from the model structure. As a result we have a language that differs from the simplest language of quantified modal logic (QS5) [19] in two points. First, all quantification is restricted by special predicates called *sorts*. We adopt the following notation proposed in [12]:

$$\text{(i) } (\forall S,x)\ A \qquad \text{(ii) } (\exists S,x)\ A$$

which can be read as *for every instance of S A holds* and *there is an instance of S such that A holds*, respectively.

In this article, (i) and (ii) are meta-linguistic abbreviations to the formulas $(\forall x\ S(x) \rightarrow A)$ and $(\exists x\ S(x) \wedge A)$, respectively, i.e., they conform to the Fregean analysis of restricted quantification. However, the primitive objects of quantification (elements of D) are continuants and as proposed in [8], the predicates used to restrict quantification represent the sortal universals that carry the principles identity, which are constitutive of the individuals that fall in their extension.

Second, individual constants of the language represent *proper names* of individuals (continuants) and, therefore, the interpretation function $\delta$ defined as

(iii) $\delta(c,w) \in D$, in which c is an individual constant
(iv) $\delta(S,w) \subseteq D$, in which S is a sort
(v) $\delta(P^n,w) \subseteq D^n$, in which P is a n-ary predicate

must obey the following constraint: for all $w,w' \in W$, $\delta(c,w) = \delta(c,w')$, i.e. the interpretation of an individual constant c (proper name) is world invariant. This amount to Kripke's thesis that proper names are rigid designators [13] and conforms to Montague's meaning postulate 1 (MP1) [8].

The quantification restricted in this way makes explicit what is only implicit in standard predicate logics. As previously discussed, suppose we want to state the following proposition : (a) *There are red tasty apples*. In classical predicate logic we would write down a *logical* formula such as (b) $\exists x\ (apple(x) \wedge tasty(x) \wedge red(x))$. In an ontological reading, (b) states that "there are things which are red, tasty and apple". The theory proposed section 2 rejects that we can conceptually grasp an individual under a general concept such as Thing or Entity or, what is almost the same, that a logic (or conceptual modeling language) should presupposed the notion of a *bare particular*. Moreover, it states that only a sortal (e.g. Apple) can carry a principle of identity for the individuals it collects, a property which is absent in attributions such as Red and Tasty. For this reason, a logical system when as used to represent a formalization of conceptual models, should not presupposed that the representations of natural general terms such as Apple, Tasty and Red stand in the same logical footing. For this reason, (a) should be represented as $(\exists Apple,x)\ (tasty(x) \wedge red(x))$ in which the sortal binding the variable x it is the one responsible for carrying its principle of identity.

Let F and G be two arbitrary universals such that F is specializes G. As a consequence we have that

1. $\Box\ (\forall F,x\ G(x))$

if G is a *rigid* universal then

2. $\Box(\forall G,x\ \Box G(x))$

or in other words, for all $w,w' \in W$ we have that $\delta(G,w) = \delta(G,w')$

For instance, Figure 1 depicts an example with the kind Person and its subkind Man. In this case we have the following instantiations of (1) and (2):

$\Box(\forall Man,x\ Person(x))$ $\Box(\forall Man,x\ \Box Man(x))$ $\Box(\forall Person,x\ \Box Person(x))$

In fact, in this example, the subkinds Man and Woman form a partition of the kind Person. In general, if $<U_1...U_n>$ is a partition of a universal U then we have that

3. $\Box(\forall U,x\ U_1(x) \oplus ... \oplus U_n(x))$

and, in this specific case, $\Box(\forall Person,x\ Man(x) \oplus Woman(x))$.

In the same figure 1, another partition is present, namely, the phase-partition Child, Adolescent, Adult of the kind Person. Phases are always defined as a partition and, thus, formula (3) always hold for a phase-partition $<K_1...K_n>$of a substance sortal S. Besides that, for all $K_a,K_b \in <K_1...K_n>$ such that $a \neq b$ we have that

4. $\Box(\forall K_a,x\ \Diamond K_b(x))$

in the example of figure 1,

$\Box(\forall Child,x\ \Diamond Adolescent(x))$ $\Box(\forall Child,x\ \Diamond Adult(x))$

$\Box(\forall Adolescent,x\ \Diamond Child(x))$ $\Box(\forall Adolescent,x\ \Diamond Adult(x))$

$\Box(\forall Adult,x\ \Diamond Child(x))$ $\Box(\forall Adult,x\ \Diamond Adolescent(x))$

Formula (4) implies

5. $\Box(\forall K_i,x\ \Diamond\neg K_i(x))$

which is a more general statement of anti-rigidity and, hence, applies to all phased-sortals including roles. In figure 2, *Student* represents a role played by instances of the kind *Person*. As previously mentioned, roles differ from phases w.r.t. their specialization conditions. In figure 1, the association enrollment $\phi_{enrollment} \subseteq Student \times School$ represents a extrinsic property that must necessary apply to all instances of Student. In general, we can state the following: Let R be a role that specializes a sortal S (named its *allowed type*) and let $\phi$ be a relation representing the restriction condition for R, such that $\phi \subseteq R \times T$, where T represents a type on which R is *externally dependent* [17]. Then,

6. $\Box(\forall R,x\ \exists T,y\ \phi(x,y))$

in the case of figure 1

$\Box(\forall Student,x\ \Diamond\neg Student(x))$ $\Box(\forall Student,x\ \exists School,y\ \phi(x,y))$

Finally, we can show why the postulate 3 (section 2) must be reinforced in conceptual models. To see that is the case suppose there is a rigid classifier G which specializes an anti-rigid classifier F. Let *{a,b,c,d}* and *{a,b}* be the extension of F and G in world *w*, respectively. By (5), there is a world *w*' in which $a \in \delta(F,w)$ is not in $\delta(F,w')$ ($a \notin \delta(F,w')$) . By (2), however, $\delta(G,w) = \delta(G,w')$ and, by (1), $\delta(G,w') \subseteq \delta(F,w')$, ergo, $a \in \delta(F,w')$ which is a contradiction. We have therefore shown that it is not the case that a rigid classifier could specialize an anti-rigid one.

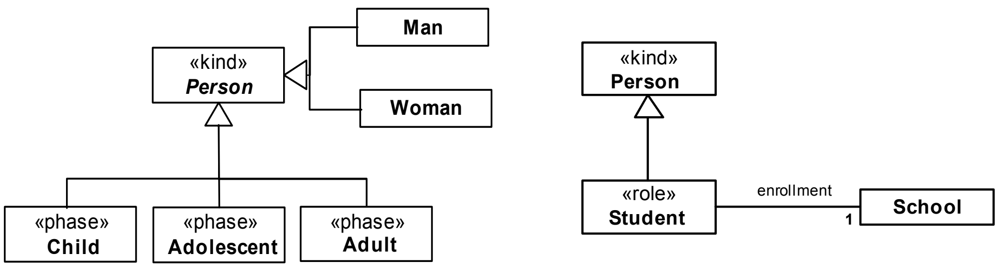


**Figure 1 (left) – Example depicting a kind and two of its partitions: a subkind-partition and a phase-partition; Figure 2 – Example depicting a phased-sortal role, its allowed type and relational restriction condition.**

## 5 Conclusions and Future Work

The development of a well-grounded, axiomatized upper level ontology is an important step towards the definition of real-world semantics for conceptual modeling diagrammatic languages. In this paper, we use a philosophically and psychologically well-founded theory of universals to address the problem of classifiers in conceptual modeling.

This theory is further used in the definition of a UML profile for Ontology Representation and Conceptual Modeling. The profile comprises of a set of stereotypes representing distinctions on types of classifiers proposed by the theory (e.g., Kind, Role, Phase, Category, Mixin) as well as a set of constraints on the possible relations to be established between these elements (representing the postulates of the theory).

A formalization of the theory is provided in a language of first-order modal logics with quantification restricted to Sortal universals. This formalization shall be extended in a future paper in which the difference between Sortals and arbitrary general terms will be emphasized. In particular, we intend to use *separated intentional properties* (in the spirit of Gupta's logic of Common Nouns [12]) to represent the intention of Sortal universals and to model the principles of identity and persistence supplied by them. This will enable us to formally address the notion of object state from an ontological point of view.

We believe that these results contribute to the task of defining ontological foundations and principled engineering tools for the discipline of conceptual modeling.

**Acknowledgements** Part of the work presented here has been developed during an extended visit of the first author to the ISTC/LOA (Trento, Italy) and in collaboration with the OntoMed research group (Leipzig, Germany). We would like to thank Nicola Guarino, Heinrich Herre, Claudio Masolo, Laure Vieu, Alessandro Oltramari, Stefano Borgo, Emanuele Botazzi, and Roberta Ferrario for fruitful discussions and for providing valuable input to the issues of this article.

## References

1. Guizzardi, G., Herre, H., Wagner G.: Towards Ontological Foundations for UML Conceptual Models. In Proceedings of 1st International Conference on Ontologies Databases and Applications of Semantics (ODBASE 2002), Springer-Verlag, Berlin, Lecture Notes in Computer Science.
2. Guizzardi, G., Herre, H., Wagner G.: On the General Ontological Foundations of Conceptual Modeling. In Proceedings of 21th International Conference on Conceptual Modeling (ER 2002). Springer-Verlag, Berlin, Lecture Notes in Computer Science.
3. Guizzardi, G., Wagner G., van Sinderen, M. An Ontologically well-Founded Profile for UML Conceptual Models, In Proceedings of the 16th International Conference on Advanced Information Systems Engineering (CAiSE), Latvia, 2004. Springer-Verlag, Berlin, Lecture Notes in Computer Science.
4. Degen, W., Heller B., Herre H., Smith, B.: GOL: Towards an axiomatized upper level ontology. In Barry Smith and Nicola Guarino, editors, Proceedings of 2nd International Conference of Formal Ontologies and Information Systems (FOIS'01), Ogunquit, Maine, USA, October 2001. ACM Press.
5. Heller, B., Herre, H. 2003. Formal Ontology and Principles of GOL. Onto-Med Report No. 1/2003. Research Group Ontologies in Medicine, University of Leipzig.
6. Object Management Group, UML 2.0 Superstructure Specification, Doc.# ptc/03-08-02, Aug. 2003.
7. Cranefield, S., Purvis M.: UML as an ontology modelling language, Proceedings of the Workshop on Intelligent Information Integration, 16th International Joint Conference on Artificial Intelligence (IJCAI-99), Germany, University of Karlsruhe (1999) 46-53.
8. van Leeuwen, J. Individuals and sortal concepts : an essay in logical descriptive metaphysics, PhD Thesis, University of Amsterdam, 1991.
9. Strawson, P. F. 1959. Individuals. An Essay in Descriptive Metaphysics. London and New York: Routledge.
10. McNamara, J. A Border Dispute, the Place of Logic in Psychology, Cambridge, M.I.T., 1986.
11. McNamara, J., Logic and Cognition, in McNamara, J.; Reyes, G. (eds.), The Logical Foundations of Cognition (Vancouver Studies in Cognitive Science, Vol 4), October 1994.
12. Gupta, A.: 1980, The Logic of Common Nouns: an investigation in quantified modal logic, Yale University Press, New Haven.
13. Kripke, S., Naming and Necessity, Harvard University Press, 1982.
14. Wiggins, D., Sameness and Substance Renewed, Cambridge University Press, 2001.
15. Perry, J. *The same F.* Philosophical Review, 1970.
16. Putnam H., Logic and Psychology, in McNamara, J.; Reyes, G. (eds.), The Logical Foundations of Cognition (Vancouver Studies in Cognitive Science, Vol 4), October 1994.
17. Guarino, N.; Welty, C., A Formal Ontology of Properties, Proceedings of the 12th International Conference on Knowledge Engineering and Knowledge Management, 2000.
18. Hirsch, E. 1982. The Concept of Identity. Oxford University Press, New York, Oxford.
19. Fitting, M; Mendelsohn, R.L., First-Order Modal Logic (Synthese Library Studies in Epistemology Logic, Methodology, and philosophy Of Science Volume 277), Kluwer Academic Publishers, 1999.