Journal of the Association for Information Systems

Volume 7 | Issue 8 Article 20

8-31-2006

Variations in Conceptual Modeling: Classification and Ontological Analysis

Irit Hadar
University of Haifa, hadari@is.haifa.ac.il

Pnina Soffer
University of Haifa, spnina@mis.hevra.haifa.ac.il

Recommended Citation
Hadar, Irit and Soffer, Pnina (2006) "Variations in Conceptual Modeling: Classification and Ontological Analysis," *Journal of the Association for Information Systems*, 7(8), .
DOI: 10.17705/1jais.00096
Available at: https://aisel.aisnet.org/jais/vol7/iss8/20

Journal of the Association for Information Systems

Special Issue

Variations in Conceptual Modeling: Classification and Ontological Analysis [1]

Irit Hadar
Department of Management Information Systems
University of Haifa
hadari@is.haifa.ac.il

Pnina Soffer
Department of Management Information Systems
University of Haifa

Abstract

Conceptual models are aimed at providing formal representations of a domain. They are mainly used for the purpose of understanding and communicating about requirements for information systems.

Conceptual modeling has acquired a large body of research dealing with the semantics of modeling constructs, with the goal to make models better vehicles for understanding and communication. However, it is commonly known that different people construct different models of a given domain although all may be similarly adequate. The premise of this paper is that variations in models reflect vagueness in the criteria for deciding how to map reality into modeling constructs. Exploring model variations as such can contribute to research that deals with the semantics of modeling constructs.

This paper reports an exploratory study in which empirically obtained model variations were qualitatively analyzed and classified into variation types. In light of the identified variation types, we analyzed two ontology-based modeling frameworks in order to evaluate their potential contribution to a reduction in variations. Our analysis suggests that such frameworks may contribute to more conclusive modeling decision making, thus reducing variations. However, since there is no complete consistency between the two frameworks, in order to reduce variations, a single framework should be systematically applied.

Keywords: Conceptual modeling, Empirical study, UML, Ontology, Variations

---

[1] Yair Wand was the accepting senior editor. There were three reviewers of whom one was Jeffery Parsons and another one was Andrew Burton-Jones. This paper was submitted on September 4 2005, and went through 2 revisions.

Introduction

Conceptual modeling, defined as “the activity of formally describing some aspects of the physical and social world around us for purposes of understanding and communication” (Mylopoulos, 1992), is applied in the early phases of information systems analysis and design. A conceptual model reflects the real world independently of implementation technology and constraints (Topi and Ramesh, 2002). It has an important role in defining, analyzing, and communicating about the requirements for the system to be. Nevertheless, it is commonly known that, given the same domain, different people may construct different models.

The immediate way of addressing differences among models is by evaluating their quality, as has been done in a number of empirical studies (Agarwal and Sinha, 1996; Batra, 1993; Kim and March, 1995; Peleg and Dori, 2000). These studies identified types of modeling errors and discussed their possible sources. Note that the evaluation of the models in these studies was based on a comparison of the models to one predefined “correct” model produced by an expert. None of the studies presented information about whether all differences were considered to be errors.

Normally, however, even those models considered “correct” may vary from each other. A question that arises is what a “correct” model is. First, it should be syntactically correct. Besides syntax, the notion of correctness in conceptual models does not have a well established and accepted definition. In general, the notion of correctness follows whatever philosophical paradigm one adopts, as discussed, for example, in Hirschheim et al. (1995). We follow Schuette and Rotthowe (1998), who referred to *construction adequacy* (rather than correctness) as a situation where consensus about the represented problem can be achieved among the people involved (e.g., designer, user). Such consensus is a subjective measure, and can in many cases be established only by directly addressing the people involved (Schuette and Rotthowe, 1998). Relating to models where such consensus can be achieved, we term differences in the elements comprising the models (e.g., entities, relationships) or in their properties (e.g., multiplicity) as model variations.

Variations among adequately constructed models may seem harmless. Nevertheless, the importance of understanding these variations is threefold. First, conceptual models serve as a basis for understanding and communicating about a problem domain. Variations may lead to difficulties and faults in such communication. Second, conceptual models are often matched for purposes such as integration (e.g. Castano, et al., 1998; Rahm and Bernstein, 2001; Palopoli et al., 2003) or reuse (Soffer and Hadar, 2003; Soffer, 2005). Third, exploring the variations among models can contribute to understanding the modeling process, and particularly the decisions made about how to map the world into modeling constructs.

Model variations and the decisions they reflect have implicitly been addressed by a large body of conceptual modeling research. Theoretical studies address the semantics of particular constructs and provide frameworks to assign semantics to a set of constructs employed by various modeling grammars. Particular constructs being addressed are part-whole relationships (e.g. Barbier et al., 2001; Barbier et al., 2003; Opdahl et al., 2001; Saksena et al., 1998; Snoeck and Dedene, 2001; Storey, 1993), associations (e.g. (Bodart et al., 2001; Evermann, 2005b; Storey, 2005; Wand et al., 1999)), and classes (e.g. Parsons, 1996; Parsons and Wand, 1997; Shanks et al., 2003). When addressing the semantics of a particular construct, guidelines are provided regarding how it can be used, and specifically what kind of real world phenomena can be expressed by which constructs. Comprehensive frameworks, such as ontologies, are aimed at providing a sound theoretical basis for conceptual modeling, including a set of constructs and the relationships among them. Such theoretical foundation is used for a variety of purposes, such as evaluating the expressive power of modeling grammars (Wand and Weber, 1993) and for analyzing specific modeling constructs and their representation of real world phenomena (e.g. Bodart et al., 2001; Opdahl, et al., 2001). In reviewing all this literature, it becomes clear that the issues investigated were motivated by difficulties and challenges raised by the researchers, not as a result of an empirical indication that such difficulties had been experienced in practice.

Empirical studies that can be related to model variations do not address variations in model creation tasks. Rather, they focus mainly on the understanding of already existing models, when different possible representation options of the domain may be considered. Specifically, such studies have investigated the implications of favoring one specific modeling construct over another. For example, Poels et al., (2005), Burton-Jones and Meso, (2002), Parsons and Cole (2004), and Shanks et al. (2005) evaluated the understanding of models, where the independent variable was the choice of a specific modeling construct to represent a real world phenomenon (where all choices are considered “correct,” or adequately structured).

However, while various aspects of model variations have been addressed separately, a broad understanding of this phenomenon has not yet been achieved. Model variations are a result of decisions made in the modeling process. These decisions, specifically, result in different people applying different ways of mapping domain phenomena into modeling constructs. A broad understanding should acknowledge difficulties incurred in modeling decisions and propose theory-based guidance to overcome such difficulties, facilitating the achievement of a higher uniformity of models.

This paper aims to take a step toward such understanding, and to this end, we address two research questions.

The first question, addressed empirically, is: Which modeling decisions may lead to model variations? Since this issue has not been extensively studied before, an exploratory study is required, to establish an understanding where no a priori hypotheses are made. Hence, we chose a qualitative research methodology in which we analyzed and classified variations among empirically obtained models. To increase external validity, we conducted the exploratory study in industry with the participation of experienced software developers.

The second research question relates to possible theory-based guidance for modeling decisions. Specifically, the guidance is needed for decisions where inconclusiveness is reflected in the empirically identified variation types. In particular, we sought guidance in the literature about the meaning and semantics of conceptual modeling constructs. Furthermore, we looked for a comprehensive framework that could potentially address a collection of modeling constructs rather than specific ones separately. This led us to the ontological interpretation of modeling constructs. Assuming that model variations reflect vagueness with respect to the semantics of modeling constructs, we expect that applying ontology-based modeling approaches will eliminate this vagueness and help in the modeling decisions. Hence, the second research question we explored in view of the empirical findings (namely, a set of identified variation types) was: Can ontological frameworks provide guidance in modeling decisions where vagueness of criteria is reflected in model variations?

We addressed this question theoretically, by analyzing two different ontological frameworks and evaluating the ability of each to provide clear guidelines where variations were found to exist. We decided to investigate more than one framework so that our conclusions would be less dependent on the specific ontology chosen. The ontological frameworks analyzed were the framework suggested by Evermann and Wand (2001a, 2001b, 2004, 2005; Evermann, 2005a) and the framework by Guizzardi et al. (2002a, 2002b, 2004). Both of these frameworks address the same modeling grammar in a coherent manner rather than each construct separately. Note that we selected these two frameworks as an example for our analysis, but other ontological frameworks coherently addressing modeling grammars can be similarly analyzed.

The remainder of this paper is organized as follows: The next section provides a discussion of model variations and their possible sources. Then we describe the empirical study, and present its findings. Next we apply two existing ontology-based modeling approaches to the variation types found in the empirical study and evaluate the potential contribution of such approaches to the reduction of model variations. We finalize with a discussion of the findings and the conclusions drawn.

Model Variations and their Sources

Soffer and Hadar (2003) proposed a framework for understanding the sources of model variations when models are constructed for a given task and purpose. Variations among models generally appear to be due to the creative nature of the modeling activity, as well as to other factors such as the richness of the spoken language (Moriarty, 2000), and the ambiguities of modeling grammars. Figure 1, which is a modification of the model presented by Topi and Ramesh (2002), presents the factors that influence the model produced by an individual for a given purpose, and their interactions. The arrows in the figure denote affecting relationships. We briefly discuss these factors and their interactions.

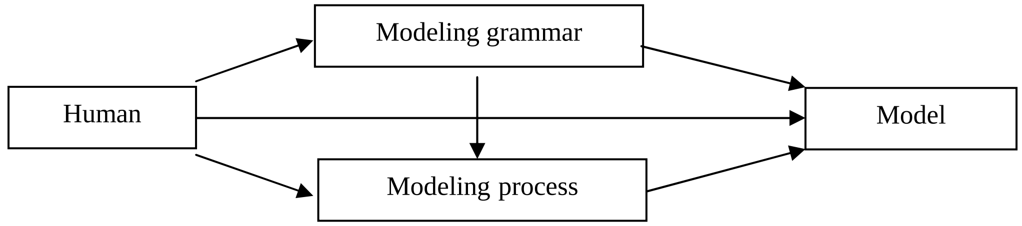


Figure 1. Factors that affect a conceptual model

Human: Human factors that may influence the model include factors affecting an individual’s perception and interpretation of reality, professional experience, and perception of model quality (e.g., some perceive a very detailed model as being of high quality, while others prefer simplicity and conciseness). Perception of reality may be a result of various factors, such as an individual’s organizational role. These perceptions may be expressed as different classifications of the domain entities into classes (Parsons, 1996) or the types of relations among them. Human factors also influence the selection and use of the modeling grammar through the modeling process and, consequently, the resulting model.

Modeling grammar: Different modeling grammars may incur different model variations due to differences in their expressive power and set of constructs involved. According to Wand and Weber (1993), expressive power means completeness (i.e., including all the constructs required for representing the domain) and clarity (i.e., without problems of construct redundancy, excess, and overload). Variations that stem from incompleteness of the grammar are due to the modelers' drive to “invent” their own solutions in order to capture more information about the real world in the model. Variations resulting from lack of clarity are caused by the fact that a single real-world phenomenon may be represented in several different ways or in an ambiguous manner.

Specific variations may occur when using modeling grammars that allow various levels of detail. In this case refinement decisions are usually made by the modeler, who applies individual judgment rather than precise rules.

Modeling process: The modeling process consists of two main phases: perceiving reality and representing it in a model. Variations can be sourced in both phases. Differences in the perception of the domain may be influenced by properties of the modelers as discussed above in relation to human factors.

The perceived reality should be mapped into modeling constructs. However, modeling grammars often employ constructs whose semantics are not precise, and they do not entail rules that conclusively define how to map real world phenomena into the modeling constructs. Hence, different mapping approaches may be taken. When modeling grammars that employ a number of views are used, different perspectives of the domain are mapped into constructs belonging to different views, and the modeling process may involve iterations among the views. The order in which these iterations are performed may also influence the mapping decisions (Shoval and Kabeli, 2001).

Once a modeling grammar is selected, the human factors that lead to potential variations are difficult to control. The modeling process is the controllable factor that can be addressed in order to reduce variations. If clear guidance to the modeling process is provided, and in particular, to the employment of the various modeling grammar constructs, variations may be reduced.

We illustrate this idea in Figure 2, which is a modified version of the model presented in Figure 1. It highlights the factors that can be controlled in order to reduce model variations. These factors are the modeling grammar, whose selection is made by the modeler (human) and determines potential variation types, and the modeling process. Control over the modeling process can be gained by applying clear modeling rules that we expect to significantly decrease the effect of human factors on the modeling process (hence, the respective arrow is dashed).

The focus of this paper is on model variations whose source is in the mapping of domain information to modeling constructs. We start by empirically exploring and categorizing model variations, and then theoretically analyzing them, specifically examining the possibility of modeling rules to reduce them.

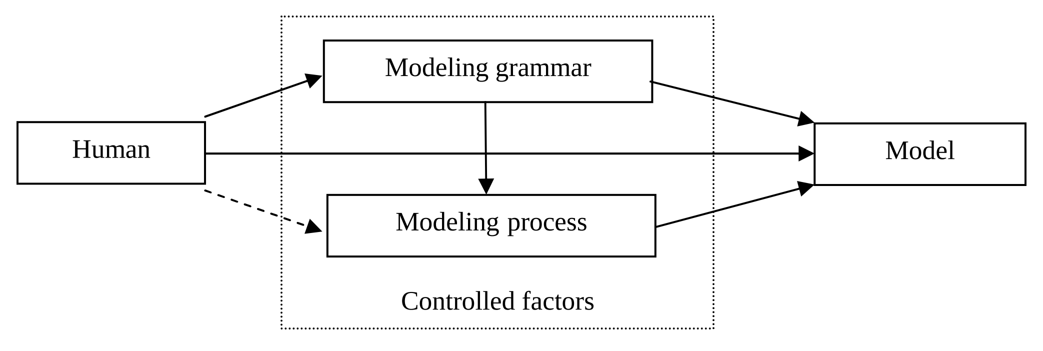


Figure 2. Controlling the factors that affect a conceptual model

Empirical Study

We designed the empirical part of this research to obtain different conceptual models of the same domain, constructed by different professional software developers. We compared the models and classified the variations among them into types. This section describes the research methodology and setting applied in the empirical study and the tasks given to the research participants. We will present the results obtained in the study, and in particular the variation types identified through the data analysis, in the next section.

Methodology and Setting

The objective of this exploratory study was to learn and understand which decisions are made during the representation of a domain in conceptual models. This can be achieved by focusing on model variations, and classifying them according to the type of decisions from which they stem.

When aiming to explore and understand a phenomenon and its different aspects, rather than statistically corroborating a hypothesis or a theory, it is appropriate to use a qualitative research approach (Bassey, 1999). In this research in particular, we used tools from the grounded theory approach (Strauss and Corbin, 1990, 1994), and we collected and analyzed data, accordingly. The empirical study was conducted with the participation of software developers employed in the software industry. The data collection was mainly based on conceptual models, created by the software developers participating in the study. In addition, we conducted individual interviews whenever the need for clarification of a written model arose. We aimed to gain an understanding of the line of thinking that led to the written model, and to examine whether a consensus could be established on whether the model was an adequate representation of the domain.

The empirical study included the following stages: (1) we gave the participants in the study a textual description of a problem domain (namely, the task), and asked them to create a conceptual model based on this description. The data collected included the written solutions (the conceptual models) created by the participants. (2) each of us separately conducted a preliminary screening of the models to evaluate the adequacy of each model. (3) we conducted follow-up interviews with modelers to further clarify modeling constructs that seemed unclear or inadequate, categorizing them as evidently wrong (and hence, not including them in the final data) when both of us found them not adequate to the problem at hand and we could not reach consensus regarding the adequacy of the model (Schutte and Rotthowe, 1998), even after conducting a clarifying interview with the modeler. In general, the models that were ruled out suffered from either logical or syntactic errors. (4) we qualitatively analyzed the data by means of identifying and coding modeling constructs and iterative classifications of model variations (further elaboration of the data analysis process is presented later in this section). Here, too, each researcher conducted the coding process separately and only then discussed the categories, in order to achieve reliability of the analysis process. We included in the final analysis only those categories on which we achieved consensus. (5) we conducted a follow up study, replicating all stages of the original one, in order to validate the results and conclusions obtained in the first study and to identify additional variation types should they appear.

The modeling grammar applied for conceptual modeling in the study was UML Class Diagram, which we chose for several reasons. First, UML is well known and considered to be a de facto standard; hence, it was relatively easy to find experienced professionals who could participate in the study. Second, it had already acquired a body of theoretical research (e.g., Evermann and Wand, 2001, 2004, 2005; Guizzardi et al., 2002, 2004), which could be of help in understanding and explaining the empirical findings. Third, we particularly selected Class Diagram as a stand-alone view, since as a first exploration of the variation phenomenon, we wanted the setting to be minimal. The rationale was to reduce the complexity so we could concentrate on understanding variations that reflect the mapping decisions using one coherent set of constructs, and avoid decisions made for keeping integrity among views. Furthermore, construct analysis in the conceptual modeling literature often addresses the constructs of Class Diagrams by themselves rather than in the context of other views (e.g., Evermann, 2005a; Barbier et al., 2003).

We selected the participants in the study according to theoretical sampling principles (Strauss and Corbin, 1990), and considered two leading factors when deciding on the sample frame. One was to ensure that the subjects practice activities related to requirements analysis, where the problem domain is analyzed and possibly modeled. The other was to make sure that the subjects are accustomed to the application of modeling as part of their job, and particularly are familiar with UML Class Diagram notation. In parallel to defining the general research population, we wished to make sure that the specific participants chosen would vary from each other in other aspects, so that the sample would not be limited by properties not relevant to the phenomenon explored. For this aim, we conducted the two empirical studies in two different firms in the information technology (IT) industry, which differ in their line of business as is further elaborated below. In addition, we checked the following aspects of the participating candidates' backgrounds: educational background, professional experience, and position in the organization (from team member to team leader to department manager). We made the final selection to ensure that the participants would represent a broad set of values with respect to these properties.

The participants in the primary empirical study were 15 software developers employed in industry, all from a company that develops information systems for production control. All the participants created and submitted conceptual models. However, after eliminating models that were found not adequate in the screening process, 11 of the 15 models were finally included in the data. The participants in the follow-up study were 14 software developers all from the same company, but different from the one in the original study.

This company develops software management and quality assurance tools for IT organizations. Here too, 11 models remained after the screening process. All participants had at least a first academic degree in computer science or software engineering (in the case of the second study – a few of the participants had obtained their degree in information systems) and professional experience of 2-12 years as software developers. Their current jobs include all phases of software development, from conceptual modeling and requirements analysis, to design, to programming and testing.

Data analysis was based on concept analysis according to the inductive analysis approach (cf. (Seidman, 1991; Strauss and Corbin, 1990; Bogdan and Biklen,1992)). In this analysis approach, categories emerge from the data and are validated and refined throughout the analysis process. For this study, we codified the variations among the obtained models for an initial categorization. The coding type applied for data analysis in this study was open coding (Strauss and Corbin, 1990), which involved breaking models down to basic model constructs (namely, classes and the relations between them), labeling each construct, comparing them, conceptualizing and, finally, categorizing model variations. The purpose of open coding analysis is to identify the emerging categories without further examination of the relationships between them. This purpose was consistent with our research objective (identifying model variation types); hence, we found open coding to be the most appropriate tool for analysis. During the iterative data analysis process, we achieved category saturation after analyzing six different models, and we found analysis of the additional five models to be consistent with the categories already identified and characterized. We then generalized this categorization into the classification presented in the following section, and re-validated against the empirical findings.

After completing the analysis of the results, we conducted the follow-up study replicating all stages of the original one. The second study took place in a different industrial environment with new participants, and a different narrative was given to the participants as the basis for conceptual modeling. These changes were made to ensure that none of the variation types identified originated from the specific characteristics of the case. Analysis of the results of the second study validated the categories and their saturation as identified in the first study. All categories identified in the original study were found in the follow-up study as well, and no new categories emerged. Since saturation had been achieved, we did not collect after the completion of the second study.

The Tasks

The task in each study was to construct a conceptual model for the purpose of understanding the problem domain, based on a given textual narrative, using UML Class Diagram. In order to avoid design considerations, the modelers were instructed not to include methods in the models. The participants in the primary study received the following textual description of a university course registration problem domain:

> A university wishes to automate its course registration procedure. The procedure is now handled as follows: For each course several course groups (also termed course offerings or sections) are available to the students. Each course group involves different lecture and lab hours, and at times, different lecturers. Students decide which courses and course groups they wish to register for, fill in a registration form and submit it to the department secretary, who manually adds the students to the relevant course groups. The planned system will enable the students to register without the secretary's involvement.

The follow-up study entailed the problem domain of a brokerage firm. The participants received the following textual description of the problem domain:

> A brokerage firm wishes to automate the clientele and stock-trade management processes applied in the business. Currently, these processes are handled as follows: The broker interacts with the stock exchange market in order to buy and sell stocks. The broker maintains a portfolio, which includes all stock holdings managed at any given moment. Clients' details are managed as well. Each client's account is managed separately, including all the stocks the client owns at a given point of time. The broker creates reports for the convenience of the clients. These reports include general stock exchange information and detailed trading activity reports. There are also defined collections of reports, where each collection provides a view of the portfolio. A client can subscribe and unsubscribe to any number of such views. However, detailed reports will include only information with regard to his/her own account.

Empirical Findings

The variations among the models produced in the study were explored and finally categorized to seven variation types. In this section the identified variation types are presented and demonstrated by using typical examples taken from the empirical data. Note that for the sake of brevity the examples presented in this paper are small and simple, focusing on demonstrating the variation types.

In addition, we present explanations based on the clarifying interviews with the modelers, as possible insights into considerations that lead to the variation types identified.

Abstract entities – Modelers may or may not use abstract entities, as illustrated in Figure 3. The *User* class in Figure 3 is abstract, i.e., it has no instances besides the instances of its specializations. In the primary study eight of the 11 models included the *User* abstract class, while three did not. Abstract entities, when used, can take various forms and sizes of inheritance hierarchies, which again increase variations. Some of the modelers used an abstract class as a generalization for any entity that is potentially a user in the system, while others modeled Professor and Student separately.

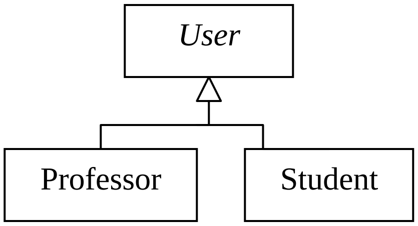


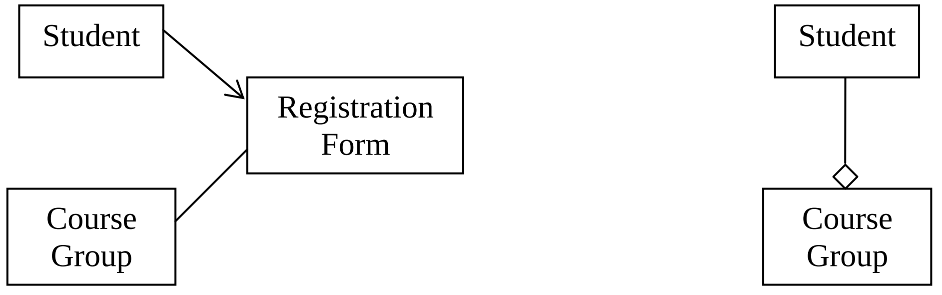
Figure 3. Abstract entities: example

We observed a similar example in the follow-up study, where in addition to an abstract *User* class (specialized to Customer and Broker), modelers defined an abstract *Account* class whose specializations are Customer Account and Broker Account. Note that there was nothing in the textual descriptions of the problem domains to suggest the inclusion of abstract classes.

Entity naming – The spoken language allows the modeler many choices of possible names. This is a common and well-recognized phenomenon (Moriarty, 2000), and was found in the present study as well. For example, the class named *User* in our previous example (Figure 3) was named *Person* by two of the modelers. These two words, although different in their semantics, hold a similar role in the models.

Variations of this type are common since there are no defined, standard, globally accepted ontologies that conclusively name domain entities. However, although different naming may influence communication and reuse application, this variation type does not affect the structure of the model as does the selection of modeling constructs. The focus of this paper is on variation types whose source is in the mapping of domain information to modeling constructs. The essence of the entity naming variation type is terminology rather than modeling construct decision. Hence, we view this variation type as being beyond the scope of this paper and it will not be investigated further here.

Class modeling decisions – These relate to the question of what elements of the real world should be modeled as classes. This issue has been addressed in general by classification theories (e.g., Lakoff, 1987), and in the context of information systems and conceptual modeling (e.g., Parsons, 1996). It has been suggested that the designation of things in the world into classes is not absolute, and may depend on culture as well as on the individual properties of the observer (modeler). In the context of this study, a modeler may face the decision of whether to classify a specific element as a class or as an association between classes. For example, in Figure 4(a) the modeler presents a Registration Form class. Figure 4(b) shows the relationship between a Student and a Course Group, where the Registration Form does not appear. Rather, the student is related to Course Group, as is the persistent situation after registration has been completed. In the primary study, three models presented Registration Form as a class, while eight models included a direct association between Student and Course Group. One participant explained that he did not included Registration Form as a class because it is a temporary thing, whose life-time is short and its essence is to create the connection between a Student and a Course Group.

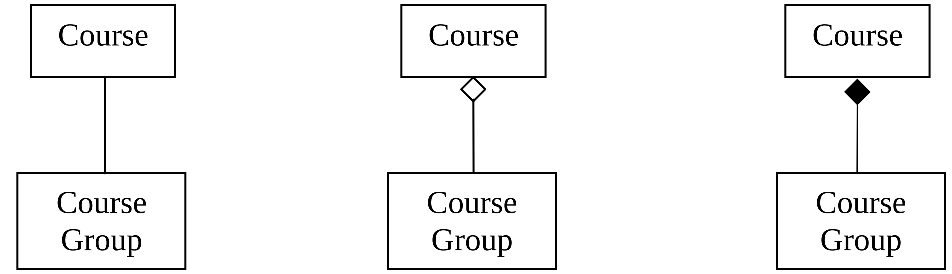


(a) Registration Form class included (b) No Registration Form class

Figure 4. Class modeling decision: example

A similar example observed in the follow-up study concerned View (collection of reports), which was sometimes modeled as a class associated to Customer and to Report, and sometimes not explicitly represented in the model, but conceptualized as an association between Customer and Report.

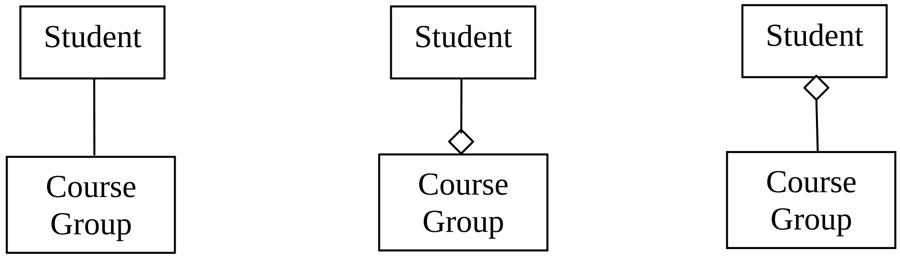
Association between classes **–** In some cases, while the existence of a relation between classes is clear, the essence of this relation is not. We found this phenomenon in cases where an aggregation or composition relation could be appropriate. For example, participants expressed the relationship between Course and Course Group classes in three different ways, as presented in Figure 5. Interestingly, the three different options appeared in similar frequencies in the data (four modeled as Composition, four as Association, and three as Aggregation). There appeared to be no solid preference for any of the association types. It seems that the modelers did not find an explicit rule to differentiate the three types of relationship. In the follow-up study, modelers also used these three relationship types (Association, Aggregation, and Composition) to express the relationship between Stock and Portfolio.



(a) Association (b) Aggregation (c) Composition

Figure 5. Different relationships between Course and Course Group Classes

Another phenomenon, related to the definition of relations between classes, appeared in the definition of the relation between Student and Course Group (Figure 6).



(a) Association (b) CG aggregates Student (c) Student aggregates CG

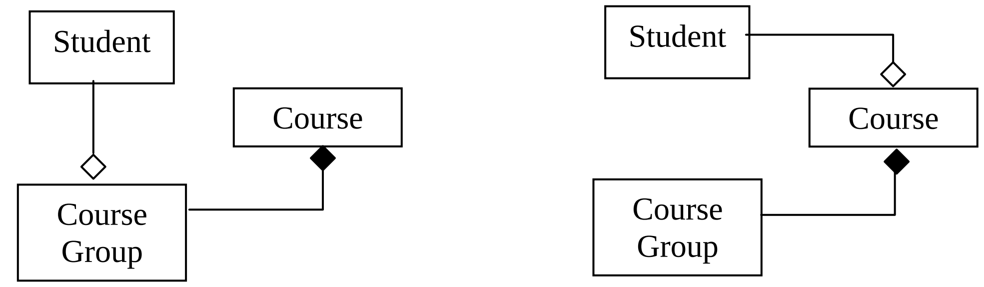
Figure 6. Different associations between Student and Course Group Classes

Here, none of the modelers thought that a Composition was in place. However, while seven of the modelers used Aggregation between Course Group and Student, four of the modelers defined an Association between them. Moreover, the modelers who presented the Aggregation relationship were divided into two similar-size groups. Four modelers presented Course Group as the whole containing the Students, while three presented the opposite whole-part direction; namely, the Student is the whole containing Course Group. In the follow-up study, we observed such inconclusiveness in the relationship between Report and View, which was modeled both as Association and as Aggregation, where Report was addressed as part of a View and vice versa.

In both tasks, this kind of inconsistency was observed only in many-to-many relationships (namely, when perceived as Aggregation, they possessed the property of shareability [2]). In this case, modelers indicated that it was more difficult to distinguish between Association and Aggregation relationships. In addition, once modelers decided on an Aggregation relationship, the many-to-many association made it difficult for them to decide which was the whole and which was the part in that relationship. In the first study, interviewed modelers mentioned the expression "has a" that was commonly used for defining Aggregation relationships. Apparently, while a Course Group "has a" Student from one perspective, it is also possible to say that a Student "has a" Course Group or a Course, from another perspective (e.g., the perspective of courses the student acquires throughout the academic path). It seems that the use of the "has a" expression in the context of Aggregation may be confusing.

Another possible cause for confusion between Aggregation and Association (but not Composition and Association) was the ability of the part to exist independently from the whole. As discussed by Barbier et al. (2001, 2003), there are nine possible cases of life-time overlapping between a part and a whole. A clear life-time dependency is easily perceived as related to a whole-part relationship, while independence may blur the nature of the relationship.

Relation of whole-part classes with other classes – In this case two classes have a whole-part relationship, and a "third party" class needs to be related to them. In such cases, it is not always clear to which of these classes, the part or the whole, the third party class should be related. Figure 7 presents an example of this variation type.



(a) Third party class associated to part (b) Third party associated to whole

Figure 7. Composite with a third party: example

Of the 11 models in the primary study, nine included the structure of Figure 7a and four included the structure of Figure 7b. Note that two models included both a relation of the Student with the whole (Course) and with the part (Course Group). Modelers who chose this conceptualization explained that while a student is part of a course group, a student is a part of a course as well. In fact, since the relation addressed here is Aggregation, modelers indicated that the transitivity of whole-part relationships contributed to this confusion. In such cases, modelers may feel it is not important whether the third party entity relates to the whole or to the part. In the case of Association, transitivity does not

---

[2] Shareability is the ability of a part to be shared by several different wholes.

play a role, however this type of confusion may still be observed. For example, in the follow-up study, this variation type appeared regarding an Association of a Broker either with Stock (part) or with Portfolio (whole).
In the above discussion, we presented the model variation types as emerged from our empirical data analysis. Table 1 summarizes the structural variation types [3] identified in the empirical study and their characterization.

Table 1. Classification of Structural Variation Types

Variation type	Description
Abstract Classes	Generalized classes that have no instance of their own. May or may not be modeled.
Association vs. Aggregation	Representation of the relation between two classes either as an Association or as an Aggregation.
Aggregation vs. Composition	Representation of the relation between two classes either as an Aggregation or as a Composition.
Directionality of Aggregation	Using different directions in an Aggregation.
Class vs. Association	Addressing a model element either as a class or as an association.
Association with Classes that have whole-part relationship	Associating the third party either to the whole or to the part class.

Examining the Potential Contribution of Ontology-Based Frameworks for Reducing Model Variations

In this section we relate to two theoretically-based frameworks that provide guidance when applying UML Class Diagrams for conceptual modeling. One framework was suggested by Evermann and Wand (2001a, 2001b, 2004, 2005) and the other by Guizzardi et al. (2002a, 2002b, 2004). We selected these two frameworks because both address UML Class Diagram in a coherent manner rather than each construct separately. We shall briefly introduce these frameworks and then evaluate their potential contribution for overcoming the variation types identified in the empirical study. Note that the aim here is not to evaluate or criticize any of these approaches. Rather, it is to examine their potential usefulness in reducing model variations, and to check whether clear modeling rules can overcome variations of all the observed types. Alternatively, types of variations might be found that are inherent in the nature of the modeling activity and cannot be eliminated by applying such an approach.

To clarify this aim, we briefly discuss different possible ways to perceive the essence of a conceptual model. Ontological frameworks may play different roles in the context of such different perceptions.

First, a conceptual model may be perceived as a faithful representation of the world. Epistemologically, this is a realist position, assuming that an objective reality exists

---

[3] We relate to structural variation types, which are variations that stem from modeling construct decisions. Particularly, the entity naming variation type is not considered as structural, hence is not included in this table.

(Casti, 1989). Taking a realist position, only a single conceptual model is acceptable as a representation of the world, and variations (which are not mistakes) cannot exist (Hirschheim et al., 1995). Using an ontological framework to achieve a single deterministic representation is logical, provided one believes the ontology being used accurately depicts reality. Hence, a person taking a realist position would not explore the possible use of different ontological frameworks, but would choose the one "truthful" ontology.

Second, a conceptual model may be viewed as a representation of how the world is perceived by an individual or a group of individuals. Epistemologically, this is a relativist position, assuming reality cannot be defined in "objective" terms, and is what a community says it is (Casti, 1989). Taking a relativist position, different representations of the world may exist, each reflecting the perception of the individual constructing the representation. A representation, once agreed upon through communication, may become shared by a group of individuals or a community. The use of an ontological modeling framework may seem to be contradictory to the relativist position, as an ontology is presumed to represent "what is in the world", i.e., assumption of an objective reality. Indeed, the use of ontologies as a basis for conceptual modeling has been criticized on these grounds (e.g., Wyssusek and Klaus, 2005). Nevertheless, while taking a relativist position, an ontological framework can still be applied for pragmatic reasons. In this case, we recognize potential difficulties that can be caused by model variations, and propose to use an ontological framework to overcome these difficulties based on semantics assigned to modeling constructs. Since a relativist position does not relate to an objective reality, different ontological frameworks can be applied without being judged on how truthfully they represent the world. They are examined here from the pragmatic perspective, namely their potential effectiveness in reducing model variations.

The Theoretical Frameworks

The Evermann and Wand Framework

The Evermann and Wand framework (2001a, 2001b, 2004, 2005) is based on Bunge's ontology (1977, 1979), as adapted for information systems modeling. Bunge's ontology relies on the philosophical foundations of Aristotle, Aquinas, Descartes and others, and presents a set of high-level, abstract constructs that are intended to be a means of representing all real-world phenomena. Evermann and Wand follow the notion of ontological expressiveness (Wand and Weber, 1993). Their fundamental premise is that in order to fully represent the world in a conceptual model, an ontological meaning should be assigned to the modeling constructs. The use of constructs without distinct ontological meaning may lead to an ontologically meaningless or ambiguous model, or to multiple model representations of the world. Their work analyzes the constructs of UML Class diagrams, State Charts, Collaboration, and Sequence diagrams, and provides rules that are intended to assure a distinct ontological meaning of these constructs. The rules include representation rules that define a mapping from the ontological constructs to the modeling constructs, and interpretation rules that map in the other direction – from UML constructs to the ontological constructs. According to Evermann (2005a), in multi-view languages such as UML, intra-diagram as well as inter-diagram modeling rules can be defined; the latter ensure the mutual integrity of different modeling perspectives (e.g., Class Diagram and State Charts (Evermann and Wand, 2005)). This paper relates only to a single view (Class Diagram); hence, the analysis relates to the intra-diagram rules addressing the constructs of Class Diagrams.

The Guizzardi et al. Framework

The Guizzardi et al. framework, similarly to Evermann and Wand, is ontology-based, and follows the notion of ontological expressiveness. The ontology underlying their work is the upper-level ontology of the General Ontological Language (GOL) (Degen et al., 2001), recently developed as a collaborative project involving philosophers, linguists and other cognitive scientists, as well as computer and information scientists. Guizzardi et al., in a series of works (2002a, 2002b, 2004), assign an ontological meaning to UML Class Diagram constructs, evaluate their expressive power, and suggest extensions to UML intended to increase its expressive power. While the analysis is ontology-based, it is also based on cognitive and linguistic theories. Unlike Evermann and Wand, who provide explicit rules for applying the UML constructs in an ontologically-meaningful way, Guizzardi et al. discuss and interpret the constructs, but do not provide explicit modeling rules in most cases. Here, we use these interpretations as potential modeling guidelines to be applied, and analyze their possible effect on the variation types found in the empirical study.

Applying the Theoretical Frameworks to the Identified Variation Types

In the following, we examine the variation types identified in the empirical study in terms of the two theoretical frameworks, and seek conclusive guidance that, when applied, would potentially eliminate variations of this type.

Abstract classes: This variation type is a result of a modeler's decision to generalize two or more classes into an abstract class, which has no instances of its own (other than the instances of its specialization classes). Both Guizzardi et al. and Evermann and Wand agree that an abstract class has no ontological meaning; that is, it is not something that exists in the real world where classes have instances. Hence, it should not be used in conceptual modeling. It can be concluded that applying each of the frameworks as guidance would eliminate this variation type.

Association vs. Aggregation: This variation type is a result of an unclear perception of what a whole-part relationship is, especially observed when the multiplicity of the relationship is many-to-many. A many-to-many multiplicity may generally exist in various types of associations. In particular, it may exist in a whole-part relationship where the parts can be shared among different wholes, namely, Aggregation.

According to Evermann and Wand, Aggregation forms a *Composite thing*, which must possess at least one *emergent* property that is not possessed by its parts but is related to them. For example, the number of registered students is an emergent property of a Course Group, in relation to its parts (Students) but not possessed by them. Applying this rule would lead to a distinction between Aggregation, where at least one emergent property (class attribute) exists in one class (the whole), and Association, where no emergent properties exist.

Guizzardi et al. disagree with this distinction and state that whole-part is the essence of structure, and should be perceived even if no emergent attribute is identified. Their claim is that the fact that an emergent attribute is not modeled does not mean that essentially there is no emergent property in the composite. While not arguing with this claim, we find it unhelpful in resolving this variation type. Nevertheless, some guidance is given by Guizzardi et al., through a discussion of transitivity that is usually associated with a part-whole relationship. For example, considering the transitivity of Aggregation relations would have helped our modelers who decided on the direction of Aggregation between Course Group and Student. Applying transitivity to the relation when it is specified in the “opposite” direction (Course Group is part of Student) shows that in the given context this direction is inappropriate. The same modelers also modeled Professor as a part of Course Group. Hence, applying transitivity in this case results in the conclusion that Professor is a part of Student, clearly an illogical conclusion. Addressing transitivity for this decision would lead to the conclusion that Student must be part of Course Group rather than the other way around.

However, the transitivity property is not always relevant to the problem domain. Neither does it provide conclusive decision rules, but rather indications. Guizzardi et al., as opposed to other researchers (e.g., Opdahl et al., 2001; Saksena et al., 1998), claim that transitivity does not always hold in part-whole relations. They suggest that in order to be meaningful and maintain transitivity, part-whole relations should be defined with respect to a *context*. A context sets the boundaries of the domain of interest, usually in terms of a granularity level. For example, a brain is a part of a person in the context of the human body, and a researcher (who is a person) is part of a research group in the context of a research organization. Due to the different contexts, it cannot be said that a brain is part of a research group.

Based on the above, transitivity can sometimes help in deciding whether to use Aggregation or Association, but this help is limited.

Aggregation vs. Composition: This variation type is common in whole-part relationships where the whole is not mandatory for the existence of the part. In such cases the part can exist without being in the composite (e.g., an engine and a car). According to Guizzardi et al., if the part cannot be shared among different wholes (e.g., an engine cannot be part of more than one car simultaneously), the relation should be modeled as Composition. This is a clear distinction that can be applied and eliminate variations of this type. Evermann and Wand claim that there is no ontological meaning to Composition, since the ontological concept of *composite* is a thing that is composed of other things that have their own existence. Hence, they suggest to avoid using Composition in conceptual modeling. Here, again, the guidance is clear, although different from that of Guizzardi et al. We will further elaborate in the discussion.

Directionality of Aggregation: This variation type involves assigning different roles to the classes participating in an Aggregation (some modelers perceive an instance of class A as being the whole whose part is an instance of class B, and others model the relationship in the opposite direction: A is part of B). This variation type did not occur with respect to Composition. As discussed previously, a possible source of this variation type is a confusion caused by the common use of the “has a” terminology. Both theoretical frameworks, while not addressing this issue directly, relate to ontologies that represent real world structures. Emphasizing structure and matter in the real world during the modeling process, the modelers will be asking themselves whether A is a part of B, or whether B is composed of A, rather than practicing “A has a B” or “B has an A.” Dealing with structural questions, such variations are not expected to occur. For example, the phrase “student has a course” sounds reasonable, while “course is part of student” does not.

Association vs. Class: This variation type is a result of the lack of a conclusive definition regarding what should and should not be modeled as a Class. In particular, the question arises when the modeled entity is not a concrete substance and is related to the interaction between two or more other entities. In such cases, the question is whether this should be modeled as a Class or as an Association.

Evermann and Wand determine that relations between things should be seen as *mutual properties* (e.g., "a person works for a company" is a mutual property of the person and the company. It is meaningless with respect to person alone or to company alone). Since properties cannot have properties of their own, they should be modeled as Associations and not as Classes. Evermann and Wand suggest representing bundles of mutual properties that relate to a single event (e.g., a purchase order) by an Association Class. They provide a set of rules to restrict the use of Association Classes and to assure that this construct is *not* confused with an “ordinary” Class which has “real” instances.

Guizzardi et al. are somewhat ambiguous with regard to this issue. According to their framework, classes represent the ontological concept of a *universal* (unit of structure), where universals can be *substance* universals but can also be *relational* universals (whose instances are relations between individuals). However, the criteria for determining whether such a universal should be modeled as a Class or as an Association are not clear. They discuss the possibility of a relational universal to have attributes and characterize situations where this can be true. Nevertheless, the guidance provided is not conclusive as to which relational universals should be modeled as Classes and which as Associations.

Association with Classes that have a Part-Whole relation between them: This variation type occurs when a “third party” Class is related to a Composite/Aggregate. The modeler faces the question of whether the Association is with the whole or with the part. Although Evermann and Wand do not explicitly address this issue, their rule with respect to Composites implies that if the relation is an emergent property of the whole (mutual with the third party), then the Association is with the whole. Otherwise, it is with the part. Guizzardi et al. do not provide guidance for this issue.

We present a summary of the findings described in this section in Table 2. The table summarizes the way each approach may contribute to the reduction of each variation type, as discussed above. It is apparent that variations of all types can be reduced by applying some kind of ontology-based modeling framework, although different frameworks may result in a reduction in variations of different types and to a different degree.

Discussion

This research aims to achieve a broad view of the model variation phenomenon as a reflection of modeling decision making. It empirically explores variation types as indicators of decision inconclusiveness, and seeks guidance in existing theoretical frameworks.

Table 2. Summary of Findings

Variation type	Evermann and Wand	Guizzardi et al.	Consistency of approaches
Abstract classes	Conclusive rule	Conclusive rule	Identical
Association vs. Aggregation	Conclusive rule	Partial guidance	Different, but not contradicting
Aggregation vs. Composition	Conclusive rule	Conclusive rule	Contradicting
Directionality of Aggregation	Implicit guidance	Implicit guidance	Identical
Association vs. class	Conclusive rule	Partial guidance	Contradicting
Association with composites	Conclusive rule	None	

The Empirical Findings

The sources of variations may relate to the individual properties of the modeler (e.g., experience), to the modeling grammar applied, and to the process of modeling (as discussed in the theoretical background). In this research, we used a specific modeling grammar (UML Class Diagram) applied by modelers with similar backgrounds in order to focus on the dependency of variations on the process of modeling.

As presented above, our assumption was that variations reflect vagueness in the criteria for deciding which construct should be used for representing the modeled domain in certain circumstances. The empirically-found variation types indicate modeling decisions for which no uniformity among modelers existed. These findings can provide a deeper understanding of the modeling process by identifying the decisions where better support and guidance could help modelers by reducing inconclusiveness.

Two limitations of the empirical study should be discussed here. First, the uniformity of domain knowledge among the subjects was not explicitly verified. The importance of domain knowledge is that having prior knowledge about the domain could influence the modeler's perception and interpretation of it (Khatri et al., 2005; Parsons and Cole, 2005). However, we used different domains in the two studies and that resulted in the same set of variation types. In the primary study it was reasonable to assume that all the modelers had a basic knowledge of the problem domain (university course registration system), since they were all university graduates. The follow-up study concerned a domain that was not related to the work experience of the modelers, although some may have been acquainted with it as customers.

Second, although the subjects of the study were experienced practitioners, their professional activities included conceptual modeling as well as software development tasks, such as design and implementation. This may have led the subjects to apply design considerations in conceptual modeling, in the study as well as in real life.

As explained in the research setting description, we conducted the selection of the study participants according to theoretical sampling principles (Strauss and Corbin, 1990). When deciding on the sample frame, we ensured that the subjects practiced activities related to requirements analysis, where the problem domain was analyzed and possibly modeled, and we made sure that the subjects were accustomed to the application of modeling techniques as part of their job, and particularly were familiar with UML Class Diagram notation. In industrial settings, the population that best met both considerations included people who were involved in a variety of tasks during the software lifecycle, from conceptual modeling and requirements analysis, to design, and sometimes even to programming and testing. Hence, design considerations observed in the study were a reflection of real life situations.

The variations addressed in this paper were found in UML Class diagrams. However, model variations are not unique to models that apply this grammar. For example, Soffer and Hadar (2003) explored variations among conceptual models that used OPM (Dori, 2002) as a modeling grammar. Comparing the findings of the current study with the ones reported for OPM, some variation types were common to both modeling grammars (e.g., the use of abstract classes), while others were grammar-specific. It may be possible to apply ontology-based sets of modeling rules to various modeling grammars, provided each set of rules is appropriate for the grammar to which it is applied. The potential effect of such rules can be analyzed with respect to observed model variations similarly to the analysis presented in this paper.

The Potential Role of Ontology-Based Frameworks

Taking a relativist position, we do not address an ontology as an absolute model of the world. Our aim was to seek guidance that would structure the modeling process and reduce the inconclusiveness that exists when mapping the real world into modeling constructs. We applied two ontology-based approaches to the variation types that we identified. In doing so, our aim was to achieve results that would not depend on a specific ontology. The results of this exercise show that using ontology-based modeling rules can indeed provide guidance in most vague situations reflected by the empirically found variations. When modeling constructs are ontologically interpreted, it becomes clear which construct represents the modeled situation as perceived by the modeler (according to the specific ontology). However, it is apparent from our results that there is no single “correct” model, as the two different frameworks provide different modeling rules. Each set of modeling rules relies on a different ontology, hence leading to a different model of the modeled domain. One may wonder which model is more “correct,” but we do not deal with this question. The question we are interested in is which framework is more useful for reducing model variations. According to Wand and Weber (2002), ultimately, the usefulness of ontologies can only be determined empirically.[4] Until an empirical evaluation of the superiority of one ontological framework over another is achieved, one may prefer to use whichever ontology one finds most helpful. In terms of guidance that may reduce variations, our findings show that Evermann and Wand’s framework provides better coverage and conclusive guidance (in terms of a clearly defined set of rules) to most of the modeling dilemmas that were reflected in the empirically-found model variations. Still, empirical evaluation of the applicability of both sets of rules is yet to be done. Hence, one may prefer the GOL ontology over Bunge’s as a representation of the world, and use Guizzardi et al.’s framework as an ontological basis. Evidently, the application of an ontology-based framework is expected to reduce variations only as long as it is used consistently, and only if the same ontology is applied by all the modelers. Consequently, an ontological basis can contribute to the

---

[4] Wand and Weber do not explicitly address how the usefulness of ontologies can be empirically evaluated. Implicitly, it is demonstrated when an ontology serves as a basis for understanding a variety of phenomena and for creating useful tools to support practical tasks.

understanding of a domain and to communicating about that domain if the underlying ontological basis is shared by all those who are involved.

We also note that ontological rules can reduce only variations whose source is in the mapping of *perceived* reality into modeling constructs. In other words, in a hypothetical situation where all modelers have the same perception of reality and all apply the same set of rules to the same grammar properly, variations are not expected to appear. However, differences in individual perceptions of the domain cannot be overcome by the use of an ontological framework, as indicated by Evermann (2005a). Considering the modeling decision process as consisting of two phases, namely, perceiving reality and mapping it to modeling constructs, an ontological framework can assist in the latter, but variations may still be caused by the former. Leaving only variations whose source is in different perceptions of the domain might help people recognize different perspectives and thereby help them learn each other’s views. The relativist position of information system design should encompass different views and perceptions of the domain (Hirschheim et al., 1995). Hence, if mapping-related variations are eliminated, the "real" perception-related variations can be highlighted, thus facilitating mutual understanding.

Conclusion

Theoretical research concerning conceptual modeling has so far been motivated by (a) the wish to provide a sound theoretical basis for a discipline that has been developed in a pragmatically-oriented manner; (b) specific observations of ill-defined semantics of certain constructs; and (c) the wish to facilitate model understanding. Model understanding has also been empirically investigated, particularly how it is affected by using specific constructs in specific situations. In relation to the construction of conceptual models, empirical studies have addressed various factors that affect the “correctness” of the produced model. It appears that the conceptual modeling research community has so far overlooked the relationship between the well-familiar variations among models and the semantics of modeling constructs as a topic for research.

One contribution of this paper is in highlighting the potential that lies in the empirical investigation of “legitimate” variations among models as an anchor to theoretical investigations of the semantics of modeling constructs.

The findings of this research are both empirical and theoretical. The empirical findings consist of a categorization of variation types that may indicate vagueness in the semantics of the constructs involved. The theoretical findings reveal that existing ontology-based frameworks can potentially contribute in all cases where such vagueness was empirically identified. Nevertheless, the findings highlight contradictions in the guidance provided by different frameworks, where differences in the underlying ontology exist. These findings emphasize the need to apply a coherent set of modeling guidelines rather than to rely on an ad-hoc collection of guidelines, separately developed for specific constructs. Such an eclectic collection may include guidelines based on different theoretical foundations. Thus, the integrity of such a collection of guidelines cannot be expected without a thorough investigation.

With respect to the possibility of reducing model variations in order to achieve better communication and to facilitate model matching and reuse operations, the findings imply that such reduction can only be achieved if all the modelers involved use the same set of rules (based on the same ontology) when constructing the model. Such reduction would only be possible with respect to variations whose source is decisions of mapping modeled domain phenomena to the constructs of the modeling technique, but not to variations caused by different perceptions of the modeled domain.

To summarize the above, while the findings emphasize the potential benefits of using ontological frameworks, they also indicate some limitations of such frameworks. These limitations include possible inconsistencies between ontologies and the uncertainty about completeness in the guidelines derived from ontologies. These limitations emphasize the need for additional research regarding these issues.

A research framework can be drawn to assess and validate existing theoretical research works, in general, and ontological frameworks, in particular. These can be validated against empirical findings of model variations, and indicate where more research is needed.

Empirical investigations are also needed to evaluate the applicability of theory-based, and in particular ontology-based, modeling rules and their actual effect on model variations. The design of such an evaluation would not be trivial. It would need to distinguish variations whose source is construct selection decisions from variations caused by the individual perception of the domain. An empirical examination of how specific rules affect each variation type in different contexts, such as different domains, may help us to understand how rules can actually be embedded in the modeling decision making, and how they interact with and influence the perception of the domain.
Another research direction is to characterize variation types with respect to different modeling grammars. The aim of such research would be to understand the inconclusive decisions that are typical of each grammar, and study the applicability of ontology-based and other sets of modeling rules with respect to different modeling grammars. This research direction can also relate to models that employ a combination of views, particularly static and dynamic ones.

Finally, this paper indicates that the choice of an ontology may affect the resulting model and that not all ontologies are equivalent in terms of modeling guidance. More research is needed to provide additional theories, possibly based on cognitive foundations, to complement the ontological frameworks and to facilitate their application.

Acknowledgement

We would like to acknowledge Yair Wand for the valuable comments that helped us focus on the main message of the paper. We would also like to thank the anonymous reviewers, whose suggestions were helpful and guided us in improving the paper.

References

Agarwal, R. and Sinha, A. P. (1996) “The Role of Prior Experience and Task Charasteristics in Object-Oriented Modeling: An Empirical Study”, *Human-Computer Studies* (45), pp. 639-667.

Barbier, F., Henderson-Sellers, B., Opdahl, A. L., and Gogolla, A. (2001) “The Whole-Part Relationship in the Unified Modeling Language: A New Approach“, In Siau,

K. and Halpin, T. (eds), *Unified Modeling Language: Systems Analysis, Design, and Development Issues*, Idea Group, pp. 185-209

Barbier, F., Henderson-Sellers, B., Parc-Lacayrelle, A., and Bruel, J. M. (2003) "Formalization of the Whole-Part Relationship in the Unified Modeling Language", *IEEE Transactions on Software Engineering* 29(5), pp. 459-470.

Bassey, M. (1999) "Chapter 7: Methods of enquiry and the conduct of case study research", in *Case Study Research in Educational Settings*, Buckingham, UK: Open University Press, pp. 65-91.

Batra, D. (1993) "A Framework for Studying Human Error Behavior in Conceptual Database Modeling", *Information and Management* (25), pp. 121-131.

Bodart, F., Patel, A., Sim, M., and Weber, R. (2001) "Should Optional Properties be Used in Conceptual Modeling? A Theory and Three Empirical Tests", *Information Systems Research* (12)4, pp. 384-405.

Bogdan, R. & Biklen, S. K. (1992) *Qualitative Research for Education: An Introduction to Theory and Methods* (2nd edition), London: Toutledge.

Bunge, M. (1977) *Treatise on Basic Philosophy: Vol. 3, Ontology I: The Furniture of the World*, Boston: Reidel.

Bunge, M. (1979) *Treatise on Basic Philosophy: Vol. 4, Ontology II: A World of Systems*, Boston: Reidel.

Burton-Jones, A. and Meso, P. (2002) "How Good are These UML Diagrams? An Empirical Test of the Wand and Weber Good Decomposition Model", in L. Applegate, R. Galliers, and J. DeGross,Eds. *Proceedings of the 23rd International Conference on Information Systems*, Barcelona, Spain.

Castano, S., De Antonellis, V., Fogini, M. G. and Pernici, B. (1998) "Conceptual Schema Analysis: Techniques and Applications", *ACM Transactions on Database System* (23)3, pp. 286-333.

Casti, J. L. (1989) "Faith, Hope, and Asperity", Ch. 1 in *Paradigms Lost*, New York: Avon Books, pp. 1-65.

Degen, W., Heller B., Herre H., Smith, B. (2001). "GOL: Towards an Axiomatized Upper Level Ontology", in Barry Smith and Nicola Guarino, (eds.) *Proceedings of FOIS'01*, USA: ACM Press.

Dori, D. (2002) *Object Process Methodology – A Holistic Systems Paradigm*, Heidelberg, New York: Springer Verlag.

Evermann, J. (2005a) "Thinking Ontologically: Conceptual vs. Design Models in UML", In Green P. and Rosemann M. (eds), *Business Systems Analysis with Ontologies,* Idea Group, Hershey, pp. 82-104.

Evermann, J. (2005b) "The Association Construct in Conceptual Modeling – an Analysis Using the Bunge Ontological Model" In Pastor O. and Cunha J. F. (eds), *Advanced Information Systems Engineering (CAiSE'05)*, (LNCS 3520), Springer-Verlag, Berlin, pp. 33-47.

Evermann, J. and Wand, Y. (2001a) "An Ontological Examination of Object Interaction in Conceptual Modeling", in *Proceedings of the Workshop on Information Technologies WITS'01*, New Orleans, pp. 91-96.

Evermann, J. and Wand, Y. (2001b) "Towards Ontologically Based Semantics for UML Constructs" in *Proceedings of ER 2001, 22nd International Conference on Conceptual Modeling,* (LNCS 2224), Berlin: Springer-Verlag, pp. 354-367.

Evermann, J. and Wand, Y. (2004) "Ontology Based Object-Oriented Domain Modeling: Fundamental Concepts", *Requirements Engineering* (10)2, pp. 146-160.

Evermann, J. and Wand, Y. (2005) "Toward Formalizing Domain Modeling Semantics in Language Syntax", *IEEE Transactions on Software Engineering* 31(1), pp. 21-37.

Guizzardi, G., Herre, H., and Wagner, G. (2002a) "Towards Ontological Foundations for

UML Conceptual Models", in *On the Move to Meaningful Internet Systems, DOA/CoopIS/ODBASE 2002 Confederated International Conferences* LNCS 2519, Berlin: Springer-Verlag, pp. 1100-1117.

Guizzardi, G., Herre, H., and Wagner, G. (2002b), "On the General Ontological Foundations of Conceptual Models", in *Conceptual Modeling - ER 2002, 21st International Conference on Conceptual Modeling,* LNCS 2505, Berlin: Springer-Verlag, pp. 65-78.

Guizzardi, G., Wagner, G, Guarino, N. and van Sinderen, M. (2004) "An Ontologically Well-Founded Profile for UML Conceptual Models" in *Advanced Information Systems Engineering, 16th International Conference, CAiSE 2004,* LNCS 3084, Berlin: Springer-Verlag, pp. 112-126.

Hirschheim, R., Heinz, K. K. and Lyytinen, K. (1995) *Information Systems Development and Data Modeling*, UK: Cambridge University Press.

Khatri, V., Ramesh, V., Vessey, I., Clay, P. and Park, S. J. (2005) "Understanding Conceptual Schemas: Exploring the Role of Application and IS Domain Knowledge", Working paper, available Feb. 2006 http://www.indiana.edu/~isdept/research/tr143.pdf.

Kim, Y. and March, S. T. (1995) "Comparing Data Modeling Formalisms", *Communications of the ACM*, (38)6, pp. 103-115.

Lakoff, G. (1987) *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*, Chicago: University of Chicago Press.

Moriarty, T. (2000) "The Importance of Names", *The Data Administration Newsletter* (15).

Mylopoulos, J. (1992) "Conceptual Modeling and Telos", Chapter 2 in P. Loucopoulos and R. Zicari (eds.), *Conceptual Modeling, Databases, and CASE*. UK: Wiley, pp. 49-68.

Opdahl, A., Henderson-Sellers, B., and Barbier, F. (2001) "Ontological Analysis of Whole-Part Relationships in OO Models", *Information and Software Technology* (43), pp. 387-399.

Palopoli, L., Sacca, D., Terracina, G., and Ursino, D. (2003) "Uniform Techniques for Deriving Similarities of Objects and Subschemes in Heterogeneous Databases", *IEEE Transactions on Knowledge and Data Engineering* (15)2, pp. 271-294.

Parsons, J. (1996) "An Information Model Based on Classification Theory", *Management Science*, (42)10, pp. 1437-1453.

Parsons, J. and Cole, L. (2004) "An Experimental Examination of Property Precedence in Conceptual Modeling", in *Proceedings of the 1st Asia-Pacific Conference on Conceptual Modeling*, Dunedin, New Zealand, pp. 101-110.

Parsons, J. and Cole, L. (2005) "What Do the Pictures Mean? Guidelines for Experimental Evaluation of Representation Fidelity in Diagrammatical Conceptual Modeling Techniques," *Data and Knowledge Engineering*, 55, 327-342.

Parsons, J. and Wand, Y. (1997) "Using Objects for Systems Analysis", *Communications of the ACM* 40(12), pp. 104-110.

Peleg, M. and Dori, D. (2000) "The Model Multiplicity Problem: Experimenting with Real-Time Specification Methods", *IEEE Transactions on Software Engineering* (26)8, pp. 742-759.

Poels, G., Gailly, F., Maes, A. and Paemeleire, R. (2005) "Object Class or Association Class? Testing the Effect on Cardinality Interpretation", in Jacky Akoka et al. (eds.), *Perspectives in Conceptual Modeling, ER'05 Workshop Proceedings* (LNCS 3770), Berlin: Springer-Verlag, pp. 33-42.

Rahm, E. and Bernstein, P. A. (2001) "A Survey of Approaches to Automatic Schema Matching", *The VLDB Journal* (10), pp. 334-350.

Saksena, M., France, R. B. and Larrondo-Petrie, M. M. (1998) "A Characteriztion of Aggregation", *Proceedings of 5th International Conference on Object Oriented Information Systems (OOIS'98)*, pp. 11-19.

Schuette, R. and Rotthowe, T. (1998) "The Guidelines of Modeling – An Approach to Enhance the Quality in Information Systems, *Proceedings of ER'98 (LNCS 1507)*, Berlin: Springer-Verlag, pp. 240-254.

Seidman, I. E. (1991) *Interviewing as Qualitative Research*. New York: Teachers College Press.

Shanks, G, Nuredini, J. and Weber, R. (2005) "Evaluating Conceptual Modelling Practices: Composites, Things, Properties", In Green P. and Rosemann M. (eds), *Business Systems Analysis with Ontologies,* Hersey, PA: Idea Group, pp. 28-55.

Shanks, G., Tansley, E. and Weber, R. (2003) "Using Ontology to Validate Conceptual Models", *Communications of the ACM* 46(10), pp. 85-89.

Shoval, P. and Kabeli, J. (2001) "FOOM: Functional- and Object-Oriented Analysis & Design of Information Systems: An Integrated Methodology", *Journal of Database Management* 12(1), pp. 15-25.

Snoeck, M. and Dedene, G. (2001) "Core Modeling Concepts in Object-Oriented Conceptual Modeling", *Proceedings of the Technology of OO Languages and Systems*, IEEE Press.

Soffer P. (2005) "Structural Equivalence in Model-Based Reuse: Overcoming Differences in Abstraction Level", *Journal of Database Management* 16(3), pp. 21-39

Soffer, P. and Hadar, I. (2003) "Reusability of Conceptual Models: The Problem of Model Variations", in *EMMSAD'03, Proceedings of the Eighth CaiSE/IFIP8.1 Int. Workshop on Evaluation of Modeling Methods in Systems Analysis and Design* (on CD).

Storey, V. C. (1993) "Understanding Semantic Relationships", *Very Large Data Bases (VLDB)* 2(4), pp. 455-488.

Storey, V. C. (2005) "Comparing Relationships in Conceptual Modeling: Mapping to Semantic Classifications", *IEEE Transactions on Knowledge and Data Engineering* 17(11), pp. 1478-1489.

Strauss, A. & Corbin, J. (1990), *Basics of Qualitative Research; Grounded Theory Procedures and Techniques*, Newbury Park: Sage.

Strauss, A. & Corbin, J. (1994) "Grounded Theory Methodology: An Overview", Chapter 1 in Denzin, N.K. and Lincoln, Y.S. (eds.), *Handbook of qualitative research,* Thousand Oaks: Sage, pp. 273-285.

Topi, H. and Ramesh, V. (2002) "Human Factors Research on Data Modeling: A Review of Prior Research, an Extended Framework and Future Research Directions", *Journal of Database Management*, (13)2, pp. 3-15.

Wand, Y., Storey, V. C. and Weber, R. (1999) "Analyzing the Meaning of a Relationship", *ACM Transactions on Database Systems* 24(4), pp. 494-528.

Wand, Y. and Weber, R. (1993) "On the Ontological Expressiveness of Information Systems Analysis and Design Grammars", *Journal of Information Systems* (3), pp. 217-237.

Wand, Y. and Weber, R. (2002) "Research Commentary: Information Systems and Conceptual Modeling – a Research Agenda", *Information Systems Research* (13)4, pp. 363-378.

Wyssusek, B and Klaus, H. (2005) "Ontological Foundations of Information Systems Analysis and Design: Extending the Scope of the Discussion", In Green P. and Rosemann M. (eds.), *Business Systems Analysis with Ontologies,* Hersey, PA: Idea Group, pp. 322-344.

About the Authors

Irit Hadar is a faculty member in the department of Management Information Systems in the University of Haifa. She has received her Ph.D. from the Department of Education in Technology and Science of the Technion – Israel Institute of Technology. Both her Master and Bachelor degrees are from the Faculty of Industrial Engineering and Management of the Technion. Her doctorate thesis examined how software development experts understand concepts and principles in Object Oriented Design. Her research areas focus on human aspects of software engineering, in particular cognitive processes, using cognitive psychology theories as analysis frameworks.

Pnina Soffer is a faculty member in the Management Information Systems department in the University of Haifa in Israel. She received her Ph.D. from the Technion – Israel Institute of Technology in 2002. In her Ph.D. thesis she developed a requirement-driven approach to the alignment of enterprise processes and an ERP system. Pnina has industrial experience as a production engineer and as an ERP consultant. Her current research areas are formal methods for business process modeling, conceptual modeling, and requirements engineering.

Journal of the Association for Information Systems

ISSN: 1536-9323

Editor
Kalle Lyytinen
Case Western Reserve University, USA

Senior Editors			
Izak Benbasat	University of British Columbia, Canada	Robert Fichman	Boston College, USA
Varun Grover	Clemson University, USA	Rudy Hirschheim	Louisiana State University, USA
Juhani Iivari	University of Oulu, Finland	Elena Karahanna	University of Georgia, USA
Robert Kauffman	University of Minnesota, USA	Yair Wand	University of British Columbia, Canada
Editorial Board			
Ritu Agarwal	University of Maryland, USA	Steve Alter	University of San Francisco, USA
Michael Barrett	University of Cambridge, UK	Cynthia Beath	University of Texas at Austin, USA
Anandhi S. Bharadwaj	Emory University, USA	Francois Bodart	University of Namur, Belgium
Marie-Claude Boudreau	University of Georgia, USA	Tung Bui	University of Hawaii, USA
Yolande E. Chan	Queen's University, Canada	Dave Chatterjee	University of Georgia, USA
Roger H. L. Chiang	University of Cincinnati, USA	Wynne Chin	University of Houston, USA
Ellen Christiaanse	University of Amsterdam, Nederland	Guy G. Gable	Queensland University of Technology, Australia
Dennis Galletta	University of Pittsburg, USA	Hitotora Higashikuni	Tokyo University of Science, Japan
Matthew R. Jones	University of Cambridge, UK	Bill Kettinger	University of South Carolina, USA
Rajiv Kohli	Colleage of William and Mary, USA	Chidambaram Laku	University of Oklahoma, USA
Ho Geun Lee	Yonsei University, Korea	Jae-Nam Lee	Korea University
Kai H. Lim	City University of Hong Kong, Hong Kong	Mats Lundeberg	Stockholm School of Economics, Sweden
Ann Majchrzak	University of Southern California, USA	Ji-Ye Mao	Remnin University, China
Anne Massey	Indiana University, USA	Emmanuel Monod	Dauphine University, France
Eric Monteiro	Norwegian University of Science and Technology, Norway	Jonathan Palmer	College of William and Mary, USA
B. Jeffrey Parsons	Memorial University of Newfoundland, Canada	Paul Palou	University of California, Riverside, USA
Yves Pigneur	HEC, Lausanne, Switzerland	Nava Pliskin	Ben-Gurion University of the Negev, Israel
Jan Pries-Heje	Copenhagen Business School, Denmark	Dewan Rajiv	University of Rochester, USA
Sudha Ram	University of Arizona, USA	Balasubramaniam Ramesh	Georgia State University, USA
Suzanne Rivard	Ecole des Hautes Etudes Commerciales, Canada	Timo Saarinen	Helsinki School of Economics, Finland
Rajiv Sabherwal	University of Missouri, St. Louis, USA	Olivia Sheng	University of Utah, USA
Ananth Srinivasan	University of Auckland, New Zealand	Katherine Stewart	University of Maryland, USA
Kar Yan Tam	University of Science and Technology, Hong Kong	Bernard C.Y. Tan	National University of Singapore, Singapore
Dov Te'eni	Tel Aviv University, Israel	Viswanath Venkatesh	University of Arkansas, USA
Richard T. Watson	University of Georgia, USA	Bruce Weber	London Business School, UK
Richard Welke	Georgia State University, USA	Youngjin Yoo	Case Western Reserve University, USA
Kevin Zhu	University of California at Irvine, USA		
Administrator			
Eph McLean	AIS, Executive Director	Georgia State University, USA	
J. Peter Tinsley	Deputy Executive Director	Association for Information Systems, USA	
Reagan Ramsower	Publisher	Baylor University	