Logics for OO information systems: a semantic study of object orientation from a categorial substructural perspective

de Haas, E.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
There is no formal transformation from the informal to the formal

(Anonymous)

Recent years have seen the convergence of many disciplines in information systems facilitated by the concepts of *Object Orientation*. Not least has been the convergence of the languages for Object Oriented analysis and design, manifested in the definition of the industrial standard UML (Unified Modeling Language [UML97], [UML99]) for such languages. Moreover the integration of information design languages into integral software development tools, enabling automatic database (persistent) model generation and code generation, indicate that these kinds of languages and concepts have grown to a mature state.

The theme of this thesis is a *semantical investigation in Object Oriented (OO) modeling and database languages*. The semantical investigation strives to give a thorough mathematical description of the concepts used in OO design and database languages. Such a mathematical description gives an insight into the constructs used, and can be used to develop and refine automatic development tools and query optimization techniques for computing with OO information objects. Fact is, most object oriented design languages, and especially UML, have no clear mathematical foundation. Nevertheless a lot of 'formal' tasks like code generation and 'database modeling' are performed in these languages. The resulting systems therefore are suspect of ambiguities and inconsistencies, and hence sometimes valid UML expressions cannot be processed. Research in the mathematical foundations of OO concepts aims to aid the development of OO language processing, by taking away the non-clarities and providing a formal and consistent way of interpreting the language.
The semantical investigation in Object Oriented design languages is especially interesting because the concepts of object orientation originate from practice and were designed to help information analysts and designers to accurately describe information models that reflect aspects of the real world. In this respect this research touches on themes from philosophy, where it is an important goal to accurately describe aspects of the real world.

In this thesis we will study the semantics of object oriented design and database languages in detail. The thesis will provide a thorough description of the concepts that can be expressed in UML and like languages We will cover all the main concepts of object orientation such as identity, inheritance, encapsulation etc.. Moreover, we will study languages for specifying information systems from a more general perspective and then identify the really basic concepts of talking about information objects. In this exercise we will encounter serious philosophical controversies that are inherent in talking about objects, but often ignored in the information system practice. It turns out that in the practice of information analysis, the information modeler runs into hard philosophical problems in his attempt to accurately describe the aspects of the real world he wants to capture.

The major artifact we will present in this thesis is a language for modeling information systems. This language contains all the main concepts of object orientation. It is a generalization of the object modeling part of UML (a fragment of the language constructs of UML, present in several diagramming techniques of UML). The basic building block of the language is a so called category and contains graphical and textual components. We will do the necessary mathematics for this language in order to obtain a formal semantics for the object oriented concepts. We will develop a formal syntactic theory for the language and provide a rigid mathematical model in which we will interpret the language. In this setting we can give a clear semantics for the basic language constructs of both object oriented modeling and design languages and object oriented database and programming languages. For the semantic study we will use the arsenal of modal and substructural logic and categorial grammars. This branch of mathematics is used heavily in the study of natural language and computation theories and the study on the OO concepts contributes a nice application of the theory with promising extensions for intelligent information systems and data mining. Moreover, we can identify the potential philosophical controversies associated with describing aspects of the real world in the information analysis practice. Such an identification will enable the information modeler to choose a consistent interpretation of the models he writes down.

Several parts of this thesis have already been communicated to the scientific community in various papers. A first version of the language of categorial graphs
This thesis is structured as follows:

- **Part 1: General analysis of Object Oriented technology.** Part 1 contains a general analysis of the concepts and intricacies of object orientation in information systems. It is the conceptualization of the domain of our semantical investigations.
  
  - In Chapter 1 we will describe the information system analysis and design practice. We will focus specifically on the object oriented analysis and design practice and the related object oriented database models. We will discuss the use of languages for analysis and design and databases, and give an overview of the languages used in practice (especially the industry standard UML). We will see that this practice imposes requirements on the language and its interpretation in the research context.
  
  - In the second chapter, we present in detail the family of notions and concepts for which we will do the semantic research. Note that much debate is possible on the exact interpretation of information system notions that originate from actual use. We will discuss the notions for object oriented (new generation) information systems in a critical way, and provide a motivation for the interpretation we will use.

- **Part 2: OO Modeling Proposal: Categorial Graphs.** In this part we propose a model in which we can research the object oriented analysis and design practice.
  
  - Chapter 3 will introduce a language for talking about information systems. This is the syntactic domain in which we can denote (graphically and textually) the concepts discussed in part 1. This language is a generalization of the common OO information system design languages. We will especially show its expressiveness by comparing it to UML. In effect, the language presented will be a formal syntactic theory for a generalized fragment of UML. The language is built from a syntactic construct we call a *categorial graph* (borrowing the term 'category' from Aristotle); and the language therefore is called the *language of categorial graphs*. 
Chapter 4 contains the semantics of the categorial graph language. We will present an interpretation of the language that talks about object oriented information systems. This interpretation will be a logic based on the theory of modal and substructural logics.

- Part 3: Logical aspects. The chapters in this part present logical aspects of the theory of object oriented information systems.

- In chapter 5 we will explain the benefits of formal semantics and describe the approach and attitude to tackle the semantics for information systems taken in this thesis. We will explain the logical aspects of doing semantics, and also position this research in the research field of logic, as it touches some very interesting problems in current logic research.

- In chapter 6 we will investigate the logic of categorial graphs. We will discuss logical aspects, especially soundness, completeness and the computational complexity of the logics for categorial graphs.

- Part 4: Philosophical backgrounds. In this part we discuss philosophical issues involved in information system modeling and object oriented concepts.

- In Chapter 7 we take a little step back, and will formulate only the basic concepts we like to have in our language that talks about information systems. We will right away discover that this basic list of desiderata already confronts us with hard problems that are (still) very actual in philosophy.

- Part 4: Conclusion This part contains a wrap up of the themes we discussed in this thesis.

- In chapter 8 we summarize what we have done and evaluate what we have achieved.