Forward Engineering and UML: From UML Static Models to Eiffel Code

Liliana Favre*, Liliana Martinez and Claudia Pereira
INTIA, Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina
* CIC (Comision de Investigaciones Cientificas de la Provincia de Buenos Aires, {lfavre, lmartine, cepereira}@cic.unicef.edu.ar

ABSTRACT

In a previous research work we have proposed a rigorous process to forward engineering UML static models. This approach is based on the integration of semi-formal notations in UML, algebraic specifications and object-oriented code. The GSBL language was defined to cope with concepts of UML models. The emphasis is given to the last steps in the road from UML to code: we describe how to transform GSBL specifications into object-oriented code. Eiffel was the language chosen to demonstrate the feasibility of our approach. In particular, we analyze the transformation of different kinds of specifications and the generation of Eiffel assertions.

INTRODUCTION

Unified Modeling Language (UML) has emerged as a standard modeling language in the object-oriented analysis and design world. It is a set of graphical and textual notations for specifying, visualizing and documenting object-oriented systems (OMG, 1999; Booch 1999).

There are CASE tools which offer code automatic generation and reverse engineering from object-oriented models. They present problems due to the lack of formal semantics of UML models and these models are semantically richer than the object-oriented languages, though.

These problems have motivated the analysis of different approaches to give semantics to the UML notations. The UML formalization is an open problem still and many research groups have already achieved the formalization of parts of the language. The Precise UML Group, pUML, is created in 1997 with the goal of giving precision to UML (Evans et al. 1998). Different results give semantics to UML subsets based on different formalisms (Lano 1995; Breu et al.1997; Brul and France, 1998; Gogolla and Ritters, 1997; Kim and Carrington, 1999; Overgaard, 1998; Barbier et. al., 2001). Currently, there are few methods that include OCL, Catalysis is the most important reference (D’Souza and Wills, 1999). Other research works propose the integration of UML with OCL (Ritters and Gogolla, 2000). Siau and Halpin (2001) and JDM (2000) identify some problematic aspects of UML and propose possible solutions.

A variety of advantages have been attributed to the use of formal software specifications to solve these problems. A formal specification can reveal gaps, ambiguities and inconsistencies. However, formal specifications alone do not address the need of industrial practitioners, who require an understandable and scalable semantics that can be integrated by using tools.

Favre and Clerici (2001) propose a rigorous process to forward engineering UML static models using the algebraic language GSBL. The contribution was towards an embedding of the code generation part in the UML model and language.

There are CASE tools for code generation starting from UML models. Unfortunately, the current techniques available in these tools are not enough for the complete automated generation of source code. As an example, Rational Rose will be analyzed (Quatrani, 1998). This allows generating databases definitions, class interfaces and relations in which the class participates. The current modeling languages available in Rational Rose (Booch, OMT, UML) do not have a precisely defined semantics. This hinders the code generation processes. Another source of problems in these processes is that, on one hand, the UML models contain information that can not be exploited in object-oriented languages and on the other, the object-oriented languages express implementation characteristics that have no counterpart in the UML models. For instance, languages like C++, Java and Eiffel do not allow explicit associations, their cardinality and their constraints. It is the responsibility of the designer to make good use of this information, selecting from a limited repertoire of implementations or implementing the association by himself. The forward and reverse processes in Rational Rose are facilitated by means of the insertion of annotations in the generated code. These annotations are the link between the model elements and the language, they should be kept intact and not changed. It is the programmer’s responsibility to know what he can modify and what he can not.

The programmer, to solve implementation problems, can modify the code by adding or removing classes, modifying class attributes or operations, changing operation signatures, etc. These code modifications require the programmer’s ability to keep the integration between model and language.

Moreover, the existing tools do not exploit all the information contained in the UML models, for instance, cardinality and constraints of associations, preconditions, postconditions and class invariants in OCL are only translated as annotations. Assertions in OCL could be translated to assertions in an object-oriented language that supports them, like Eiffel. Moreover, reuse is based on object-oriented language libraries and not on specifications that describe relations between classes and their operations free from implementation details.

To overcome these problems, a rigorous process to forward engineering UML static models using the algebraic language GSBL was proposed (Favre and Clerici, 2001).

Starting from UML class diagrams, an incomplete algebraic specification can be automatically built by instantiating reusable schemes.

This paper appears in Issues and Trends of Information Technology Management in Contemporary Organizations, the proceedings of the Information Resources Management Association International Conference.
This specification contains the highest information obtained by translating the UML constructions and OCL constraints to GSBL\textsuperscript{oo}. Preconditions, postconditions and invariants in OCL will be translated to GSBL\textsuperscript{oo}. The process of transformation is supported by a transformation rules system (Favre et al., 2000).

Transformations are based on a reusable component library defined by the SprReIm model. (Favre and Clérici, 2001). This model allows defining structured collections of reusable components that integrate algebraic specifications and object-oriented code. It takes advantage of the algebraic formalism power to describe behavior in an abstract way integrating them with concrete implementations. It consists of three abstraction levels:

- **Specialization**: It describes a hierarchy of incomplete specifications related by specialization relationships through two views: one based on OCL specifications and the other on GSBL. Preconditions, postconditions and invariants in OCL will be translated to GSBL\textsuperscript{oo}
- **Realization**: It describes a hierarchy of complete specifications related by realization relationships.
- **Implementation**: It relates hierarchy of concrete classes schemes in an object-oriented language.

The reuse of components is based on the application of operators. Reuse operators were defined on the three levels of the SprReIm model.

Thus, an algebraic specification can be semi-automatically built. It can be used to detect inconsistencies in the class diagrams. The GSBL\textsuperscript{oo} specifications must be integrated with object-oriented code. The emphasis in this paper is given to the transformation of GSBL\textsuperscript{oo} to Eiffel.

### FORMALIZING UML STATIC MODELS IN THE GSBL\textsuperscript{oo} LANGUAGE

The existing algebraic specification languages do not possess specific constructions to express all kinds of relations in UML (dependency, association, generalization and realization). These are generally buried in client and inheritance relations. However, associations are semantics constructions of equal weight to the classes and generalizations in the UML models and should not be treated just as implementation constructions. In fact, the associations allow abstracting the interaction between classes in the design on large systems and they affect the partition of the systems in modules. Since an integrated method requires common constructing mechanisms for object-oriented models and algebraic specifications, the GSBL\textsuperscript{oo} language has been defined. It enriches GSBL (Clérici, 1988). The ability to specify deferred and effective parts, inheritance relations among classes and mechanisms such as implicit parameterization to support reuse of specifications and their incremental development are among the most important features of GSBL.

GSBL\textsuperscript{oo} extends GSBL with constructions that allow expressing different kinds of UML relations. OBJECT CLASS and ASSOCIATION class specify classes and associations (ordinary, qualified, association-class) respectively.

An OBJECT CLASS distinguishes incomplete and complete parts. The DEFERRED clause declares new sorts, operations or equations incompletely defined. The EFFECTIVE clause either declares new sorts, operations or equations completely defined, or completes the definition of some inherited sort or operation. Sorts and operations are declared in the SORTS and OPS clauses. In GSBL\textsuperscript{oo}, it is possible to specify partial functions restricting operations by preconditions.

GSBL\textsuperscript{oo} expresses dependencies in the context of classes by means of the USES clause. The mechanisms for defining inheritance are the RESTRICT and REFINES clauses. The REFINES clause relies on the module viewpoint of classes and the RESTRICTS clause on the type viewpoint of a class. Both of them reflect IS-A relations between abstractions in the external model of any software system. Associations are defined as standard elements in GSBL\textsuperscript{oo}.

Generic relations can be used in the definition of concrete relations by instantiation. New associations and whole-part relations (aggregation and composition) can be defined by means of the following syntax:

**ASSOCIATION**

\[
\text{ASSOCIATION} \langle\text{relationName}\rangle
\]

\[
\text{IS} \langle\text{constructorTypeName}\rangle\left[...:\text{Class1};...:\text{Role1};...:\text{Role2};...:\text{visibility1};...:\text{visibility2}\right]
\]

\[
\text{CONSTRAINED BY} \langle\text{constraintList}\rangle
\]

**WHOLE-PART**

\[
\text{WHOLE-PART} \langle\text{relationName}\rangle
\]

\[
\text{IS} \langle\text{constructorTypeName}\rangle\left[...:\text{Whole};...:\text{Class1};...:\text{Role1};...:\text{Role2};...:\text{visibility1};...:\text{visibility2}\right]
\]

\[
\text{CONSTRAINED BY} \langle\text{constraintList}\rangle
\]

The IS clause expresses the instantiation of \langle\text{constructorTypeName}\rangle with classes, roles, visibility and multiplicity. The CONSTRAINED-BY clause allows the specification of static constraints in first order logic. Relations are defined in an Object Class by means of the following syntax:

**OBJECT CLASS**

\[
\text{OBJECT CLASS} \langle\text{c}\rangle
\]

\[
*:\text{ASSOCIATES} \langle\text{className}\rangle
\]

\[
*:\text{HAS-A SHARED} \langle\text{className}\rangle
\]

\[
*:\text{HAS-A NON-SHARED} \langle\text{className}\rangle
\]

... 

**END-CLASS**

The keywords ASSOCIATES and HAS-A identify ordinary association or aggregation respectively. The keywords SHARED and NON-SHARED refer to simple aggregation and composition respectively. An association may be refined to have its own set of operations and properties. Such an association is called an Association Class.

The PACKAGE is the mechanism provided by GSBL\textsuperscript{oo} for grouping classes, and controls its visibility. It matches the UML semantics. Classes and their relations from a system design into a series of packages might be separated using the GSBL\textsuperscript{oo} import dependencies to control access among these packages.

Figure 1 shows an information system for a company. There is an association between Person and Company, specifying that managers (instances of Person) manage companies. Every manager may manage only one department and every company may have only one manager. There is a qualified association between Person and Bank. In the context of Bank, an accountNumber (qualifier) identifies a particular customer. In an employer/employee relation between Company and Person, there is a Job that represents the properties of that relation, which applies to exactly one pairing of Person and Company. Figure 2 partially shows the GSBL\textsuperscript{oo} specification of Figure 1.

Figure 1: Company information system. A UML class diagram.
Figure 2: Information system company. A GSBL specification

TRANSFORMING GSBL<sup>oo</sup> SPECIFICATIONS IN EIFFEL CODE

The resulting specifications of transforming UML static models in GSBL<sup>oo</sup> must be integrated with an object-oriented code. These transformations are described below and they are exemplified for the classes and relationships expressed in the UML diagram in Figure 1.

Transformation of Classes
The algebraic specification obtained in the previous step must be transformed into Eiffel code. To achieve this, every clause and relation present in a GSBL<sup>oo</sup> OBJECT CLASS specification was analyzed. GSBL<sup>oo</sup> and Eiffel have the same syntax for the declaration of class parameters. Then, this transformation is reduced to a trivial translation.

The relation introduced in GSBL<sup>oo</sup> using the clause USES will be translated into a client relation in Eiffel. The relation expressed through the keywords REFINES and RESTRICTS in GSBL<sup>oo</sup> will become an inheritance relation in Eiffel. This provides mechanisms to carry out modifications on the inherited classes that will allow adapting them.

The DEFFERRED and EFFECTIVE clauses in GSBL<sup>oo</sup> declare sorts and operations of the class with the equations that define their behavior. If an OBJECT CLASS is incomplete, i.e., it contains sorts and operations in the clause DEFFERED, the keyword class in Eiffel is preceded by the keyword deferred. Sorts do not require explicit translation.

From the signature of the operations, the interfaces for the methods of the Eiffel class are generated (feature in Eiffel). The translation of each operation has a different treatment according to the type of feature to which it makes reference (functions, procedures, variables or constants). It should also be considered that of all the domains of an operation, the first one that coincides with the sort of the specified class is the object Current in the object-oriented language and it should be eliminated from the list of parameters of the resulting feature.

Functions and procedures can present arguments. Once each name is obtained, either by an explicit requirement to the user or by extracting it from the specification, the list of arguments of each feature is built. Functions and procedures require a body defined by the keywords do end, which will be completed by the user.

Transformation of Axioms
Eiffel provides an assertion language. Assertions are boolean expressions expressing semantics properties of the classes. They “serve to express the specification of software components: indications of what a component does rather than how it does it” (Meyer 97). They can play the following roles:

• Precondition: expresses the requirements the client must satisfy to call a routine.
• Postcondition: expresses the conditions the routine guarantees on return.
• Class invariant: expresses the requirements every object of the class must satisfy after its creation.

Preconditions and axioms of a function written in GSBL<sup>oo</sup> are used to generate preconditions and postconds for routines and invariants for Eiffel classes.

It is worth clarifying for the assertion generation that a basic functionality \( f: x \times a_1 \times a_2 \times \ldots \times a_n \), where \( x \) is the sort of interest, is translated into Eiffel syntax as \( f(a_1,a_2,\ldots,a_n) \).

A GSBL<sup>oo</sup> precondition, which is a well-formed term defined over functions and constants of the global environment classes, is automatically translated to Eiffel method precondition.

Axioms in a formal specification language represent the constraints the class introduces on the operations. Axioms are translated to Eiffel postconditions and invariants.

The system can automatically derive an invariant if it can establish a correspondence between the functions in the axiom and the class features that only depend on the object state.
A postcondition can automatically be generated from one axiom if a term $e(<\text{list-of-arguments}>)$, associated to an operation, can be distinguished within itself in such a way that any other term of the axiom depends upon the $<\text{list-of-arguments}>$ or constants. Then, the postcondition will associate itself with the method that reflects the term and will obviously depend only upon the previous state of the method execution, upon the state after its execution and upon the method arguments.

Another type of situation cannot be automatically derived. These cases do not enable the system to build an assertion without the user's interventions. The programmer can also incorporate assertions that reflect purely implementation aspects. A detailed description and examples may be found in Favre (1998).

**Transformation of Associations**

Associations are transformed by instantiating reusable schemes that express how to implement them. These schemes are linked to the different relation constructor types in GSBL.

The transformation of associations is automatically derived starting from existent schemes in the level of implementation of the Association component.

The specialization level describes the different relations through incomplete specifications classified according to its kind, degree and connectivity. The realization level describes a hierarchy of specifications associated to different realizations. For instance, for an association (binary, bi-directional and many-to-many) different realizations through hashing, sequences or trees could be associated. The implementation level associates each leaf of the realization level to different implementations in an object-oriented language.

Implementation sub-components express how to implement associations and aggregations. For example, a bi-directional binary association with multiplicity “one-to-one” will be implemented as an attribute in each associated class containing a reference to the related object. On the other hand, if the association is “many-to-many” the best approach is to implement the association as a different class, in which each instance represents one link and its attributes.

For each HAS-A clause, an implementation scheme will be selected and the “aggregate” and the “part” will be instantiated. For example, if the aggregation is “one-to-many”, for an attribute in the “aggregate”, a reference to a sequence of pointers to the “part” will be generated.

Analogously, for every ASSOCIATES clause, a scheme in the implementation level of the Association component will be selected and instantiated. In these cases, the implementation level schemes suggest including “reference” attributes in the classes or introducing an intermediate class or container. Notice that the transformation of an association does not necessarily imply the existence of an associated class in the generated code, as an efficient implementation can suggest including “reference” attributes in the involved classes.

Figure 3 partially shows the Eiffel resulting code of transforming the association with multiplicity “one-to-one”.

**REFERENCES**


Related Content

Developing Appreciative College Experience with Personal Learning Networks
www.irma-international.org/chapter/developing-appreciative-college-experience-with-personal-learning-networks/112793

Enhancing Formal Methods with Feature Models in MDD
www.irma-international.org/chapter/enhancing-formal-methods-with-feature-models-in-mdd/112409

An Efficient Random Valued Impulse Noise Suppression Technique Using Artificial Neural Network and Non-Local Mean Filter

A Particle Swarm Optimization Approach to Fuzzy Case-based Reasoning in the Framework of Collaborative Filtering

Analysis by Long Walk: Some Approaches to the Synthesis of Multiple Sources of Evidence
www.irma-international.org/chapter/analysis-long-walk/28263