Formal Specification and Analysis of Domain Specific Languages using Maude

José E. Rivera, Francisco Durán, and Antonio Vallecillo
Dpto. de Lenguajes y Ciencias de la Computación
Universidad de Málaga, Spain
{rivera,duran,av}@lcc.uma.es

Abstract

Modeling languages play a cornerstone role in Model-Driven Software Development for representing models and metamodels. Modeling languages are usually defined in terms of their abstract and concrete syntax. This allows the rapid development of languages and some associated tools (e.g., editors), but does not allow the representation of their behavioral semantics, something especially important in certain industrial environments in which simulation and verification are critical issues. In this paper we explore the use of Maude as a formal notation for describing models, metamodels, and their dynamic behavior, making models amenable to formal analysis, reasoning and simulation.

1 Introduction

Domain-Specific Modeling (DSM) is a way of designing and developing systems that involves the systematic use of Domain Specific Languages (DSLs) to represent the various facets of a system, in terms of models. Such languages tend to support higher-level abstractions than general-purpose modeling languages, and are closer to the problem domain than to the implementation domain. Thus, a DSL follows the domain abstractions and semantics, allowing modelers to perceive themselves as working directly with domain concepts. Furthermore, the rules of the domain can be included into the language as constraints, disallowing the specification of illegal or incorrect models.

DSLs play a cornerstone role in DSM. So far, defining a modeling language involved at least two aspects: the domain concepts and rules (abstract syntax), and the notation used to represent these concepts (concrete syntax)—let it be textual or graphical. The abstract syntax of a DSL is usually defined by a metamodel. A metamodel describes the concepts of the language, the relationships between them, and the structuring rules that constrain the model elements and their combinations in order to respect the domain rules. This approach enables the rapid and effective development of DSLs and some associated tools (e.g., editors, constraint checkers or metric evaluators).

Current DSM approaches have mainly focused on the syntactic (i.e., structural) aspects of DSLs. Explicit and precise specification of the behavioral semantics of models has not received much attention by the DSM community until recently, despite the fact that this creates a possibility for semantic mismatch between design models and modeling languages of analysis tools. While this problem exists in virtually every domain where DSLs are used, it is more common in domains in which behavior needs to be explicitly represented. This issue is particularly important in safety-critical real-time and embedded system domains, where semantic ambiguities may produce conflicting results across different tools. Furthermore, the lack of explicit behavioral semantics strongly hampers the development of formal analysis and simulation tools.

In order to remedy this situation, a growing number of proposals is trying to express the behavioral semantics of metamodels using diverse approaches (see Section 6). However, most of them present several limitations, especially when it comes to represent complex behaviors at a high level of abstraction. Moreover, many of them do not count with proper tool support for the formal analysis of the specification produced.

In this paper we explore the use of Maude [11, 12] as a formal notation and system for supporting the specification of modeling languages. In particular, we show how models, metamodels and their dynamic behavior can be specified in Maude using rewriting logic, and how we can make use of its analysis capabilities to reason about them: the fact that rewriting logic specifications are executable allows us to apply a flexible range of increasingly stronger formal analysis methods and tools, such as reachability analysis [12], model checking [21], or theorem proving [13]. Maude also offers a comprehensive toolkit for automating such kinds of formal analysis of specifications, efficient enough to be of practical use, and easy to integrate with software development environments such as Eclipse (http://www.eclipse.org/).
The structure of this document is as follows. First, Section 2 introduces the importance of adding semantics to a metamodel. Then, Section 3 serves as a brief introduction to Maude. Section 4 describes how models, metamodels and their dynamic behavior can be represented in Maude. Section 5 is dedicated to show how this Maude representation makes models amenable to formal analysis, providing some application examples. Finally, Section 6 compares our work with other related proposals and Section 7 draws some conclusions and outlines some future research activities.

2 Models, Metamodels and their Semantics

A model is written in the language of its metamodel. A metamodel describes the concepts of the language, the relationships between them, and the structuring rules that constraint the model elements and their combinations in order to respect the domain rules. We normally say that a model conforms to its metamodel.

Metamodels are also models, and therefore they need to be written in another language, which is described by its meta-metamodel. This recursive definition normally ends at the meta-meta-level, since meta-metamodels conform to themselves.

As described by Chen et al. in [10], the semantics of a DSL may be either structural or behavioral. The structural semantics describe the meaning of the models in terms of the structure of model instances: all of the possible sets of components and their relationships, which are consistent with the well-formedness rules, are defined by the abstract syntax. The behavioral semantics describe the evolution of the state of the modeled artifacts along some time model. Hence, the behavioral semantics need to be formally captured by a mathematical framework representing the appropriate form of dynamics.

In general, behavioral semantics of metamodels have not been explicitly specified, or this has been done using natural languages [26]. This way of describing semantics is not precise enough, and therefore it may lead to several (mis)interpretations of the same model. When a model is used to represent some ideas, illustrate a design or just as a roadmap to the corresponding implementation, semantics may be not needed. But this is not the case in DSM.

Some authors point out that it is normally possible to guess the meaning of most terms of a metamodel, since a good language designer probably chooses keywords and special symbols with a meaning similar to some accepted norm. But a computer cannot act on such assumptions [24]. To be useful in the computing arena, any language (either textual or visual, either used for programming, requirements, specification or design) must come with rigid rules that clearly state allowable syntactic expressions and give a rigid description of their meaning. This issue is especially important for realizing some model operations like simulation and verification, and is crucial in the safety-critical real-time and embedded systems domain, where semantic ambiguities may produce conflicting results across different tools.

Several notations have been proposed to specify models and metamodels. However, the kind of formal expressiveness and tool support they provide is quite limited (see Section 6). In this paper we propose Maude as a formal notation for describing models, metamodels and their dynamic behavior, providing metamodels with semantics. Contrary to other proposals where abstract syntax and semantics live at different worlds, we can specify both abstract syntax and semantics using Maude, with no need of another notation for specifying the mappings. Furthermore, these specifications will allow models to be simulated and analyzed by the Maude system and its formal analysis environment [13].

3 Rewriting Logic and Maude

Membership equational logic (MEL) [7] is a Horn logic whose atomic sentences are equalities \( t = t' \) and membership assertions of the form \( t : S \), stating that a term \( t \) has sort \( S \). Such a logic extends order-sorted equational logic, and supports sorts, subset relations, subset polymorphic overloading of operators, and definition of partial functions with equationally defined domains. MEL has sound and complete rules of deduction, and initial and free algebras [7].

Rewriting logic (RL) [30] is a logic of change that can naturally deal with state and with highly nondeterministic concurrent computations. In RL, the state space of a distributed system is specified as an algebraic data type in terms of an equational specification \( (\Sigma, E) \), where \( \Sigma \) is a signature of sorts (types) and operations, and \( E \) is a set of equational axioms. The dynamics of a system in RL is then specified by rewrite rules of the form \( t \rightarrow t' \), where \( t \) and \( t' \) are \( \Sigma \)-terms. These rules describe the local, concurrent transitions possible in the system, i.e. when a part of the system state fits the pattern \( t \) then it can change to a new local state fitting pattern \( t' \). The model theory of RL has been developed in detail in [30], where initiality, soundness and completeness theorems are proved.

In what follows, we will use the syntax of Maude [12], a wide spectrum programming language directly based on RL, with MEL as its underlying equational theory, to present rewrite theories. Thus, Maude integrates an equational style of functional programming with RL computation. Because of its efficient rewriting engine and its metalinguage capabilities, Maude turns out to be an excellent tool to create executable environments for different logics, models of computation, theorem provers, or even programming languages. In addition, Maude has already been successfully used in software engineering tools and several ap-
plications [29]. We informally describe in this section those Maude’s features necessary for understanding the paper; the interested reader is referred to [12] for details.

For example, the following Maude functional module natural numbers (with sorts Nat of natural numbers and NzNat of nonzero natural numbers), using the Peano notation, with the zero (0) and successor (s__) operators as constructors (note the ctor attribute). The addition operation (___+) is also defined, being its behavior specified by two equational axioms. The operators __s__ and __+_ are defined using mixfix syntax (underscores indicate placeholders for arguments).

```
**Figure 1. A production system metamodel.**
```

Computation in a functional module is accomplished by using equations as simplification rules from left to right until a canonical form is found. Some equations, like those expressing the commutativity of binary operators, are not terminating but nonetheless they are supported by means of operator attributes, so that Maude performs simplification modulo the equational theories provided by such attributes, which can be associative (assoc), commutativity (comm) and identity (id).

While functional modules specify membership equational theories, rewrite theories are specified by system modules. A system module may have the same declarations of a functional module plus rules. Unconditional rules are written with syntax `rl [l] : t => t ′`, and conditional rules as `crl [l] : t => t ′ if Cond ′`, with `l` the rule label and `Cond` its condition. The guards of conditional rules act as blocking pre-conditions, in the sense that a conditional rule can only be fired if its condition is satisfied.

Assuming a parameterized module SET specifying sets of elements, with a sort Set{X} of sets, empty as the empty set, and ___. the associative and commutative union operator, with identity element empty, let us consider the following BOARD system module.

```
The rewrite command can be used to execute the system, using the Maude interpreter, which applies the rules and stops when no rule can be applied.

Maude> rewrite 6, 3, 2 .
result NzNat: 4
```

Note that the final result depends on the numbers selected in each application of the rule.

4 Models, Metamodels and their Semantics in Maude

In [39] we presented a proposal for representing and manipulating models based on the use of Maude, which not only was expressive enough for these purposes, but also offered good tool support for operating with models. In particular, we showed how some basic operations on models, such as model subtyping, type inference, and metric evaluation, can be easily specified and implemented in Maude, and made available in development environments such as Eclipse. That work is extended here to cope with the specification of the behavioral semantics of metamodels, something we already outlined in our previous work [38]. In this paper, that representation has been improved for efficiency reasons. Moreover, here we show how to make use of Maude’s formal analysis methods and tools to reason about the model specifications.

4.1 A Production System Example

For illustration purposes, let us introduce a modeling language for industrial production lines, which will serve as the motivating example to show the capabilities of our approach. It is inspired by a common example used in several works for describing DSLs (see, e.g., [9, 43]).

The production systems metamodel is shown in Fig. 1. A production system is composed of different kinds of machines: generators, transfers, assemblers and containers. Generators produce items, transfer machines move them through different trays, assemblers consume them to create
new ones, and containers store the assembled items. Machines take their inputs from and put their results in trays which contain items up to their capacity.

Fig. 2 shows a production model for hammers. It is composed of generators of heads and handles, a transfer machine that moves the generated heads and handles to the input tray of an assembler, an assembler that combines heads and handles to form hammers, and a container that stores (an unlimited number of) hammers. The figure also shows the number of handles and heads to be created (attribute counter of class Generator) as well as the maximum capacity of trays.

In the following, we will refer to the model in Fig. 2 as initModel (the initial model). Another model, resulting from modifying both generator counters model to the value −1 (making generators create items indefinitely) will be considered as well. We will refer to this second model as infInitModel.

### 4.2 Representing Models with Maude

Although the notion of model is not yet fully agreed, we like adopting the definition given by Jean Bézivin and Frédéric Jouault in [4]: a model $\mathcal{M}$ is a triplet $\mathcal{M} = (G, \omega, \mu)$ where $G$ is a directed multigraph, $\omega$ is itself a model (called the reference model or metamodel of $\mathcal{M}$) associated to a graph $G_\omega$, and $\mu$ is a function associating elements (nodes and edges) of $G$ to nodes of $G_\omega$. The relation between a model and its reference model is called conformance.

Models are represented by sets of Maude objects. Nodes are represented as Maude objects, and node attributes by object attributes. Edges are represented by object attributes too, each one representing the reference (by means of object identifiers) to the target node of the edge. Thus, models are structures of the form $\text{mm}\{\text{obj}_1, \text{obj}_2, \ldots, \text{obj}_N\}$, where $\text{mm}$ is the name of the metamodel, and the $\text{obj}_i$ are the objects that represent the corresponding nodes. Maude objects are record-like structures of the form $<o : c | a_1 : v_1, \ldots, a_n : v_n>$, where $o$ is the object identifier, $c$ is the class the object belongs to, $a_i$ are attribute identifiers and $v_i$ their corresponding current values.

Given the appropriate definitions for all classes, attributes and references in its corresponding metamodel (as we shall see in next section), the following Maude term is a possible representation of the hammers production model depicted in Fig. 2.

```
ProductionSystem {
  < 'heg : HeadGen | counter : 10,
      in : empty, out : 't1 >
  < 'hag : HandleGen | counter : 10,
      in : empty, out : 't2 >
  < 't1 : Tray | items : empty, in : 'heg,
      out : 'tf, capacity : 2 >
  < 't2 : Tray | items : empty, in : 'hag,
      out : 'tf, capacity : 2 >
  < 'tf : Transfer | in : ('t1, 't2), out : 't3 >
  < 't3 : Tray | items : empty, in : 'tf,
      out : 'as, capacity : 4 >
  < 'as : Assembler | in : 't3, out : 't4 >
  < 't4 : Tray | items : empty, in : 'as,
      out : 'co, capacity : 2 >
  < 'co : Container | in : 't4, out : null,
      items : empty > }
```

Note that quoted identifiers (identifiers with a quote in front of them) are used as object identifiers. Relations with multiplicity * are represented as sets of references (of sort Set(Old)). Relations with arity 0..1 are represented as values of sort Maybe(Old), which can be either an object identifier or the constant null.

Figure 2. A Production System Model.
4.3 Representing Metamodels with Maude

As part of the infrastructure we provide, there is a sort for every metamodel element: a sort @Class for classes, a sort @Attribute for attributes, a sort @Reference for references, etc. Thus, a metamodel is represented by declaring a constant of the corresponding sort, for each metamodel element. More precisely, each class is represented by a constant of a sort named after the class. This sort, which will be declared as subsort of sort @Class, is defined to support class inheritance through Maude’s order-sorted type structures. Attributes and references are represented by constants of subsorts of @Attribute or @Reference, respectively, depending on their cardinality, and on whether their elements have to be ordered or can be duplicated. For instance, sort @AttributeSimple represents attributes with a [1-1] multiplicity; sort @ReferenceMaybe represents references with a [0-1] multiplicity; and sort @ReferenceSet represents non-ordered collections of non-duplicated references.

To illustrate this approach, the following fragment of Maude specification describes the production system metamodel depicted in Fig. 1.

```maude
mod PRODUCTION-SYSTEM is
  extending MAUDELING-BASE.
  op ProductionSystem : -> @Metamodel.
  op PS : -> @Package.
  sort Machine.
  subsort Machine < @Class.
  op in : -> @ReferenceSet.
  op out : -> @ReferenceMaybe.

  sorts Assembler Transfer.
  subsorts Assembler Transfer < Machine.
  op Assembler : -> Assembler.
  op Transfer : -> Transfer.

  sort Container.
  subsort Container < Machine.
  op Container : -> Container.
  op items : -> @ReferenceSet.

  sort Generator.
  subsort Generator < Machine.
  op Generator : -> Generator.
  op counter : -> @AttributeSimple.

  sort HeadGen HandleGen.
  subsorts HeadGen HandleGen < Generator.
  op HeadGen : -> HeadGen.
  op HandleGen : -> HandleGen.

  sort Item.
  subsort Item < @Class.
  op Item : -> Item.

  sort Head Handle Hammer.
  subsorts Head Handle Hammer < Item.
  op Head : -> Head.
  op Handle : -> Handle.
  op Hammer : -> Hammer.
  sort Tray.
  subsort Tray < @Class.
  op Tray : -> Tray.
  op capacity : -> @AttributeSimple.
  op items : -> @ReferenceSet.
  op in : -> @ReferenceSet.
  op out : -> @ReferenceMaybe.
endm
```

Other properties of metamodel elements, such as whether a class is abstract or not, the opposite of a reference (to represent bidirectional associations), or attributes and reference types, are expressed by means of Maude equations defined over the constant that represents the corresponding metamodel element. We can find, for example, the following definitions:

```maude
*** class Tray
  eq isAbstract(Tray) = false.
  ...
*** attribute capacity (of class Tray)
  eq type(capacity) = @Int.
  eq lowerBound(capacity) = 1.
  eq upperBound(capacity) = 1.
  ...
*** reference items (of class Tray)
  eq type(items) = Item.
  eq opposite(items) = null.
  eq lowerBound(items) = 0.
  eq upperBound(items) = *.
  ...
```

Classes, attributes and references are qualified with their containers’ names, so that classes with the same name belonging to different packages, as well as attributes and references of different classes, are distinguished. For example, assuming that all the classes in the model are included in a package PS, identifiers Tray, capacity and in would really be Tray@PS, capacity@Tray@PS and in@Tray@PS (or in@Machine@PS). These qualifications are omitted here to improve readability.

Note that this notation allows users to get all the metamodel information from a model with no need to access to the Maude metalevel. Thus, operations that need metamodel information are not only easier to specify, but also more efficient. A good example is the operation for validating models, which checks whether a model conforms to its metamodel (verifying, among other conditions, cardinalities and type constraints defined for a model).

Finally, please notice that metamodels are models too, and thus they can also be represented by set of objects. The classes of such objects will be the ones specified in the meta-metamodels, for example, the classes that define the MOF (or Ecore) metamodel. In this way, metamodels can be handled in the same way as models are, i.e., operations and analysis techniques defined over models can be applied to metamodels too.
4.4 Representing Behavior with Maude

As we have previously mentioned, explicit and formal specification of behavioral semantics has not received much attention by the DSM community so far, despite the fact that this lack opens the door to the possibility for semantic mismatches.

For instance, it is not clear from the production system metamodel when are handles and heads generated, or what is the precise behavior of the system when one of the trays is full. Can we execute the system for simulating different trays capacities, in order to find the most efficient production line? These are the sort of issues that need to be precisely clarified by a behavioral specification.

Dynamic behavior can be specified in Maude in terms of rewrite rules added to the corresponding metamodel specification. Thus, given a Maude module with the specification of a metamodel, we can “extend” it with behavioral information (by adding the appropriate rewrite rules).

To illustrate this approach, the following Maude module PRODUCTION-SYSTEM-WITH-BEHAVIOR extends the above PRODUCTION-SYSTEM module with five rules.

\[
\text{mod PRODUCTION-SYSTEM-WITH-BEHAVIOR is}
\]
\[
\text{inc PRODUCTION-SYSTEM .}
\]
\[
\text{vars HEG T HE HA IT A HAM CO T' T'' TR : Oid .}
\]
\[
\text{vars ITEMS ITEMS' ITEMS'' : Set\{Oid\} .}
\]
\[
\text{vars SFS SFS' SFS'' SFS''' SFS'''' : Set\{@StructuralFeatureInstance\} .}
\]
\[
\text{var OBJSET : Set\{@Object\} .}
\]
\[
\text{vars CONT CAP CAP' CAP'' : Int .}
\]
\[
\text{crl [GenHead] :}
\]
\[
\text{ProductionSystem{}
\]
\[
< \text{HEG : HeadGen | counter : CONT, out : T, SFS >}
\]
\[
< \text{T : Tray | capacity : CAP, items : ITEMS, SFS' >}
\]
\[
\text{OBJSET } \}
\]
\[
\Rightarrow
\]
\[
\text{ProductionSystem{}
\]
\[
< \text{HEG : HeadGen | counter : CONT - 1, out : T, SFS >}
\]
\[
< \text{T : Tray | capacity : CAP, items : ITEMS', SFS' >}
\]
\[
< \text{HE : Head |}
\]
\[
\text{OBJSET } \}
\]
\[
\text{if HE := headId(CONT)}/
\]
\[
\text{\ / CONT /= 0} /
\]
\[
\text{\ / ITEMS' < CAP .}
\]
\[
--- \text{GenHandle rule similar to GenHead}
\]
\[
\text{crl [Transfer] :}
\]
\[
\text{ProductionSystem{}
\]
\[
< \text{T : Tray | items : (IT, ITEMS), out : TR, SFS >}
\]
\[
< \text{T' : Tray | items : ITEMS', out : TR, SFS' >}
\]
\[
< \text{TR : Transfer | out : T'', SFS'' >}
\]
\[
< \text{T'' : Tray | capacity : CAP', items : ITEMS'', SFS''' >}
\]
\[
\text{OBJSET } \}
\]
\[
\Rightarrow
\]
\[
\text{ProductionSystem{}
\]
\[
< \text{T : Tray | items : ITEMS, out : A, SFS' >}
\]
\[
< \text{A : Assembler | out : T', SFS''' >}
\]
\[
< \text{T' : Tray | capacity : CAP', items : ITEMS', SFS''' >}
\]
\[
\text{OBJSET } \}
\]
\[
\text{if | ITEMS' | < CAP' .}
\]
\[
\text{crl [Assemble] :}
\]
\[
\text{ProductionSystem{}
\]
\[
< \text{HA : Handle | SFS }>
\]
\[
< \text{HE : Head | SFS' }>
\]
\[
< \text{T : Tray | items : (HE, HA, ITEMS), out : A, SFS'' >}
\]
\[
< \text{A : Assembler | out : T'', SFS'''' >}
\]
\[
< \text{T'' : Tray | capacity : CAP'', items : ITEMS'', SFS'''' >}
\]
\[
\text{OBJSET } \}
\]
\[
\Rightarrow
\]
\[
\text{ProductionSystem{}
\]
\[
< \text{T : Tray | items : ITEMS, out : A, SFS' >}
\]
\[
< \text{A : Assembler | out : T', SFS''' >}
\]
\[
< \text{T' : Tray | capacity : CAP', items : ITEMS', SFS''' >}
\]
\[
\text{OBJSET } \}
\]
\[
\text{if HAM := hammerId(HE, HA)}/
\]
\[
\text{\ / | ITEMS' | < CAP' .}
\]
\[
\text{rl [Store] :}
\]
\[
\text{ProductionSystem{}
\]
\[
< \text{T : Tray | items : (IT, ITEMS), out : CO, SFS >}
\]
\[
< \text{CO : Container | items : ITEMS', SFS'' >}
\]
\[
\text{OBJSET } \}
\]
\[
\Rightarrow
\]
\[
\text{ProductionSystem{}
\]
\[
< \text{T : Tray | items : ITEMS, out : CO, SFS' >}
\]
\[
< \text{CO : Container | items : (IT, ITEMS'), SFS'' >}
\]
\[
\text{OBJSET } \} .
\]
\text{endm}

The GenHead and GenHandle rules specify the behavior of head and handle generators, respectively (the latter is not shown because it is analogous to the former). Since the value of the tray attribute counter is decreased every time an item is generated, a head (or handle) is created provided that: (a) there is room for it in the out tray (the number of items in it is smaller than its capacity), and (b) the number of items already created is smaller than its upper bound (i.e., its attribute counter is different from zero). The headId (resp. handleId) function returns a new identifier ‘hen’ (resp. ‘han’) given a natural number n. The \(|_\_| \) function returns the cardinality of a given set.

The Transfer rule specifies the behavior of transfer
machines. If there is an item in one of its in trays, it will place it in its out tray provided that there is enough space in it.

The Assemble rule specifies the behavior of assemblers. If there is an assembler with a head and a handle in its in tray, and space on its out tray, then it will consume such a head and a handle and will put a hammer in its out tray. Given identifiers ‘hen’ and ‘ham’, the hammerId function returns a new identifier ‘hen.ham’.

The Store rule represents the hammer storage. Every time a hammer is placed on the a container’s in tray, the container stores it.

Note that the dynamic semantics of the model is precisely specified by these rewrite rules. For instance, if the out tray of a generator is full, no further items are created (until some item is taken from it); in case that a transfer can move several items from its in trays to its out one, only one (at a time) will be moved in an nondeterministic way; etc.

Moreover, this specification can be used for simulation and execution purposes, as well as for formal analysis. We will see below how this nondeterminism can be removed by using Maude’s strategies (Section 5.1) for controlling the execution of the system.

Finally, please note that several alternative behaviors can be specified by defining other modules extending the original PRODUCTION-SYSTEM module with different rules. This, together with the module inheritance Maude supports, provides a very powerful mechanism for refining and extending behavioral specifications.

5 Formal Analysis

Once the system specifications are written using the modeling approach presented in the previous sections, what we get is a rewriting logic specification of the system. Since the rewriting logic specifications produced are executable, this specification can be used as a prototype of the system, which allows us to simulate and analyze it.

Maude offers tool support for interesting possibilities such as model simulation, reachability analysis and model checking. We focus here on these three facilities, which will be described with the help of examples that illustrate Maude’s possibilities in this context. Detailed descriptions of the commands and tools used here can be found in [12]. In addition, models can be further analyzed using other available tools in Maude’s formal environment: the LTL satisfiability and tautology checker, the Church-Rosser checker, the coherence checker, the termination tool, etc. We refer the interested reader to [12, 13] for details on these tools.

5.1 Simulation

Given a Maude specification as the one described in the previous sections, it can be executed by just successively applying equations and rewrite rules on an initial term (model). Maude provides two different rewrite commands, namely rewrite and frewrite, which implement two different execution strategies, a top-down rule-fair strategy, and a depth-first position-fair strategy, respectively [12].

The rewrite and frewrite commands explore a possible execution of different initial models. However, a rewrite system do not need to be Church-Rosser and terminating, and there might be many different execution paths. Although these commands are enough in many practical situations where an execution path is sufficient for testing executability, the user might be interested in exploring all possible execution paths from the starting model, a subset of these, or a specific one.

There are currently three ways of controlling the execution process in Maude, which, ordered from the more complex to the easiest to use, are:

- Using the reflective capabilities of Maude, with built-in metalevel functions such as metaReduce, metaApply, etc., which provide absolute control over the execution process [12].
- Using the Maude strategy language [20] to define strategy expressions that control the way terms are rewritten.
- Using invariant-driven strategies for executing specifications complying with given invariants [18].

We illustrate here the use of the strategy language as a mechanism to guide the execution process. In this language, a strategy is described as an operation that, when applied to a given term, produces a set of terms as a result (since in general this is a nondeterministic process). The basic strategies consist of the application of a rule (identified by the corresponding rule label) to a given term, and allowing variables in a rule to be instantiated before its application with a specific substitution. Basic strategies are combined by means of operators such as concatenation (\(|\)\), union (\(|\)\), iteration (+ for zero or more times and \(*\) for one or more times), if-then-else, etc.

We can, for example, use the Maude srewrite_using command to find all possible applications of the rules specified in Section 4.4 from initModel as follows:

```
srewrite initModel using all .
```

\(^{1}\)For MEL specifications, being Church-Rosser and terminating means not only confluence—so that a unique normal form will be reached—but also a sort decreasingness property, namely that the normal form will have the least possible sort among those of all other equivalent terms.
In this case the number of possible executions is huge, and although may be helpful for some automatic checking, the result is not very useful here. More interesting is the result given by the following command:

```
srewrite initModel
    using (GenHead ; Transfer ; GenHandle ; Transfer ; Assemble ; Store) + .
```

In this case, the following sequence is enforced: firstly, a head is generated and moved to the in conveyor of the assembler; then, a handle is generated and moved as well; afterwards, both the head and the handle are assembled to generate a hammer; finally, this hammer is stored in the container. Ten solutions are found, one for each possible time that the sequence can be repeated before the generators reach their limit. For instance, the last solution found, where ten hammers are created and stored, is the following:

```
ProductionSystem{
    < 'as : Assembler | in : 't3, out : 't4 >
    < 'co : Container | in : 't4, out : null,
        items : (O1,O2,O3,O4), SFS >
    < 'h1 : Hammer | empty >
    < 'h2 : Hammer | empty >
    < 'h3 : Hammer | empty >
    < 'h4 : Hammer | empty >
    ...
    < 'h10 : Hammer | empty >
    < 'hag : HandleGen | counter : 0,
        in : empty, out : 't2 >
    < 't1 : Tray | capacity : 2, items : empty,
        in : 'hag, out : 'tf >
    < 't2 : Tray | capacity : 2, items : empty,
        in : 't1, out : 'tf >
    < 't3 : Tray | capacity : 4, items : empty,
        in : 't2, out : 'as >
    < 't4 : Tray | capacity : 2, items : empty,
        in : 'as, out : 'co >
    < 'tf : Transfer | in : ('t1, 't2), out : 't3 >
}
```

The number of solutions to be shown can be limited using the corresponding (optional) argument of the `srewrite` command. Thus, for example, we can ask for the first fifteen executions, using a less strict strategy (applying some of the rules with some freedom), as follows:

```
srewrite [15] initModel
    using ((GenHead ; Transfer) | (GenHandle ; Transfer) | Assemble | Store) + .
```

Among the solutions obtained we get the one above (together with many others). The `rewrite` and `frewrite` commands also allow users to specify an upper bound for the number of rule applications. This upper bound can be very useful to simulate non terminating system, or to execute a simulation step by step.

### 5.2 Reachability Analysis

Executing the system using the `rewrite` and `frewrite` commands means exploring just one possible behavior of the system. As in Section 5.1, this can be partially improved by using rewriting strategies. However, it does not allow us to analyze the solutions found. The Maude `search` command allows us to explore (following a breadth-first strategy up to a specified bound) the reachable state space in different ways.

Let us start by looking for deadlock states. Although in this example we may easily guess how the deadlock states look like, let us be a bit naive and look for final states (states on which no further rewrite may take place) using the `!!` variant of the `search` command, with no hammers in the final container. For instance, starting from our initial model, and given variables `OBJSET` of sort `Set[@Object]`, we can use the `search` command as follows:

```
search [10] initModel =>!
ProductionSystem{
    < 'co : Container | items : empty, SFS >
}
```

Please note that we have limited the number of solutions to 10, otherwise it gives hundreds of solutions. By observing them we realize that a possible source of deadlock is that the `t3` tray may be full of items of the same type, not allowing the assembler machine to proceed. Let us look now for states satisfying this condition. For instance, starting from our initial model, and given variables `O1`, `O2`, `O3` and `O4` of sort `Id`, `I` of sort `Item`, `SFS` of sort `Set[@StructuralFeatureInstance]` and `OBJSET` of sort `Set[@Object]`, we can search for states on which the tray `t3` contains four items of the same type, using the `search` command as follows:

```
search [1] initModel =>*
ProductionSystem{
    < 't3 : Tray | items : (O1,O2,O3,O4), SFS >
    < O1 : I | empty >
    < O2 : I | empty >
    < O3 : I | empty >
    < O4 : I | empty >
    OBJSET }
```

Note the use of the `=>*` variant of the `search` command to look for any state, final or not. Note as well that we have requested only the first solution. The result given by Maude is the following:

```
Solution 1 (state 210)
OBJSET -->
< 'as : Assembler | in : 't3, out : 't4 >
< 'co : Container | in : 't4, out : null,
    items : empty >
```
As a result, a collection of solutions (variable substitutions) that fulfill the given search pattern is obtained. In this case, a state with four heads placed in the tray ‘t3’ is found.

The shortest sequence of rewrites followed to reach this model can be displayed by typing show path n, where n is the number of a state. If we are only interested in the labels of the applied rules, the command show path n can be used instead. In this case, show path labels 210 produces the sequence GenHead GenHead Transfer GenHead Transfer GenHead Transfer Transfer, which is one of the possible paths leading to this state.

Upper bounds for the number of rule applications and solutions can also be specified in the search. For instance, let us consider now the infInitModel initial model (with the counters of generators set to −1). Since the counter attributes are decreased each time an item is generated, and generators stop when its value is equals to zero, it can result in a nonterminating execution if hammers are correctly produced (containers do not have an upper bound for contained items). Then, it could be interesting to know, for example, whether a model on which the container stores at least ten hammers can be reached. Given variables ITEMS of sort Set{Oid}, SFS of sort Set{@StructuralFeatureInstance} and OBJSET of sort Set{@Object}, we can check whether starting from initModel the capacity of any tray is exceeded:

```
search initModel =>*
ProductionSystem {
  < O : Tray |
    capacity : CAP, items : ITEMS, SFS > OBJSET }
  such that | ITEMS | > CAP .
```

No solution.

Since no solutions are found, we can state that (starting from initModel) the capacities of the trays are never exceeded.

Whenever an invariant is violated, we are guaranteed to get a counterexample (up to time and memory limitations). However, when the number of states reachable from the initial state is infinite, and the invariant holds, we will search forever, never finding it. In these cases, bounded search (a search with a maximum number of rule applications) is a widely used procedure that: (a) together with a big depth can greatly increase the confidence that the invariants hold, and (b) can be quite effective in finding counterexamples, where applicable.

We can do the same check for infInitModel with a bound on the depth of the search of 1000 as follows.

```
search [1,1000] infInitModel =>*
ProductionSystem {
  < O : Tray |
    capacity : CAP, items : ITEMS, SFS > OBJSET }
  such that | ITEMS | > CAP .
```

No solution.

With this information, we could very easily modify the dynamic behavior of our model to avoid deadlocks. For example, we could add a condition to the Transfer rule so that an item is never put in the output tray if all the items in it are of the same type (of the item).

5.3 LTL Model checking

Maude offers a linear temporal logic explicit-state model checker [21], which allows us to check whether every possible behavior starting from a given initial model satisfies a given temporal logic property. Maude’s model checker can be used to prove safety and liveness properties of rewriting systems when the set of states reachable from an initial state is finite. Full verification of invariants in infinite-state system can be accomplished by verifying the invariants on finite-state abstractions [31] of the original infinite-state system, that is, on an appropriate quotient of the original system whose set of reachable states is finite [12].
In order to specify safety and liveness properties, state predicates are needed. For instance, suppose that we want to check that a hammer that has been created is eventually stored in the container. In this case, we need two different state predicates: exist and stored, that will reflect whether an item exists or is stored, respectively.

State predicates are defined as operators of sort Prop, and their semantics are given by means of a set of equations that specify for what model pattern a given state predicate evaluates to true. In Maude, opposite to the standard LTL propositional logic, we can define parametric state predicates, that is, operators of sort Prop which need not be constants, but may have one or more sorts as parameter arguments. We include the following declarations in a module including the MODEL-CHECKER module:

\[
\begin{align*}
\text{ops} & \quad \text{exist} \quad \text{stored} \quad \colon \text{Oid} \to \text{Prop} . \\
\text{var} & \quad \text{O} \quad \text{IT} \quad \colon \text{Oid} . \\
\text{var} & \quad \text{ITEMS} \quad \colon \text{Set\{Oid\}} . \\
\text{var} & \quad \text{SFS} \quad \colon \text{Set\{@StructuralFeatureInstance\}} . \\
\text{var} & \quad \text{IC} \quad \colon \text{Item} . \\
\text{var} & \quad \text{OBJSET} \quad \colon \text{Set\{@Object\}} . \\
\text{var} & \quad \text{M} \quad \colon \text{@Model} . \\
\text{var} & \quad \text{P} \quad \colon \text{Prop} . \\
\text{eq} & \quad \text{ProductionSystem} \{ < \text{O} : \text{IC} \mid \text{SFS} > \text{OBJSET} \} = \text{exist}(\text{O}) = \text{true} . \\
\text{eq} & \quad \text{ProductionSystem} \{ < \text{O} : \text{Container} \mid \text{items} : (\text{IT}, \text{ITEMS}), \text{SFS} > \text{OBJSET} \} = \text{stored}(\text{IT}) = \text{true} . \\
\text{eq} & \quad \text{M} = \text{P} = \text{false} \quad \text{[otherwise]} . \\
\end{align*}
\]

After defining these state predicates, we are ready to model-check different LTL properties. Making use of the eventually $<>$ and henceforth $[]$ LTL connectives, we can model check, for instance, that a created hammer (‘H10.S10) is stored at some time, by typing:

\[
\text{reduce modelCheck(initModel, } [[] \text{exist('he10.ha10) -> <> stored('he10.ha10))} . \\
\text{result Bool: true}
\]

Of course, this property is not very useful, but it becomes a bit more interesting if we check it for all possible hammers. Given the exist and stored propositions above, the following formula allHammers expresses the same property for all possible hammers, with identifiers created as in the above rules:

\[
\begin{align*}
\text{vars} & \quad \text{N} \quad \text{M} \quad \text{K} \quad \colon \text{Nat} . \\
\text{op} & \quad \text{allHammers} \quad : \text{Nat} \to \text{Formula} . \\
\text{op} & \quad \text{allHammers} \quad : \text{Nat} \to \text{Nat} \to \text{Formula} . \\
\text{eq} & \quad \text{allHammers}(\text{N}, \text{M}) = \text{allHammers}(\text{N}, \text{M}, \text{M}) . \\
\text{eq} & \quad \text{allHammers}(\text{s N}, \text{0}, \text{K}) = [[] \text{exist(ID) -> <> stored(ID))} \\
& \quad \land \text{allHammers(N, K, K)} . \\
\text{if} & \quad \text{ID} := \text{hammerId(headdId(s N), handleId(s M))} . \\
\text{eq} & \quad \text{allHammers}(\text{0, 0}, \text{K}) = [[] \text{exist(ID) -> <> stored(ID))} \\
& \quad \land \text{if} \quad \text{ID} := \text{hammerId(headdId(0), handleId(0))} . \\
\end{align*}
\]

We can now check it as follows:

\[
\text{reduce modelCheck(initModel, allHammers(10,10))} . \\
\text{result Bool: true}
\]

Of course, not all properties are true. For example, if we model-check that a created head, say ‘he10, is consumed at any time (to form a hammer) by typing:

\[
\text{reduce modelCheck(initModel, } []\![\text{exist('he10) -> <> ~ exist('he10))} . \\
\text{resulrt Bool: true}
\]

We get a counterexample showing why it fails (in this case due to deadlock states).

### 5.4 Tool Support

One of the main advantages of using Maude is due to its execution environment, able to provide efficient implementations of the specifications—comparable in resource consumption to most commercial programming languages’ environments.

Our work so far has consisted in providing support inside Eclipse for specifying models and metamodels, and their behavior using Maude, and to perform the analyses described above [36].

Our Maudeling tool and all the specifications in this paper are available at [http://atenea.lcc.uma.es/](http://atenea.lcc.uma.es/).

### 6 Related work

One way to specify the semantics of a language is to define a translation from expressions in that language into expressions in another language with well defined semantics [24]. These semantic mappings between semantic domains are very useful not only to provide metamodels with semantics but also to be able to simulate, analyze or reason about them using the logical and semantical framework available in the target domain [14] (in general, each semantic domain is more appropriate to represent and reason about certain properties, and to conduct certain kinds of analyses).

For example, some works specify semantics by using visual rules [16, 19, 43], which prescribe the preconditions of the actions to be triggered and the effects of such actions. These pre- and postconditions are given visually as models that use the concrete syntax of the DSL. This kind of representation is quite intuitive, because it allows designers...
to work with domain specific concepts and their concrete syntax for describing the rules [17]. Graph transformation is one of these rule-based approaches. The most common formalization of graph transformation is the so-called algebraic approach, which uses category theory to express the rewriting [19]. This approach supports a number of interesting analysis techniques, such as detecting rule dependencies [19] or calculating critical pairs (minimal context of pairs of conflicting rules) [25].

However, graph transformation offers limited support for other kinds of analyses, such as reachability analysis, model checking, etc. This is why some authors define semantic mappings between graph transformation and other semantic domains, and then back-annotate the analysis results to the source notation [9, 16, 17, 25]. This possibility allows one to use the techniques specific to the target semantic domain for analyzing the source models. For example, in [2] rules are translated into Alloy in order to study the applicability of sequences of rules and the reachability of models; in [1] rules are translated into Petri nets to check safety properties; in [44] they are transformed into Promela for model-checking; and in [8, 9] rules are transformed into OCL pre- and postconditions for rule analysis (e.g., conflict detection) using standard OCL tools.

One of the problems of these approaches is due to the fact that they require, from the DSL designer, deep knowledge of the target language in order to specify the transformations. However, this problem can be partially overcome if the transformations can be automated, using, e.g., model transformation techniques. For example, in [17] the authors are able to generate the transformations from rule-based Domain-Specific Visual Languages (DSVLs) into semantic domains with an explicit notion of transition, such as Place-Transition Petri Nets. The generated transformation is expressed in the form of operational triple graph grammar rules that transform the static information (initial model) and the dynamics (source rules and their execution control structure). Similarly, in [9] the authors describe how graph transformation rules can be mapped into OCL pre- and postconditions for rule analysis. However, as the own authors describe in their paper, not all kinds of graph transformations can be automatically transformed into OCL, and the analyses that can be conducted are somehow limited.

Other approaches use UML behavioral models to represent the system dynamics. For example, in [22] operational semantics are represented using UML collaboration diagrams, which are then formalized into graph transformation rules. In [23], Story Diagrams are presented as a new graph rewrite language based on UML and Java. More precisely, the authors propose the use of UML together with chunks of Java code to express a graph rewrite language mixed with object-oriented data concepts. The whole specification (including dynamic behavior) is then transformed to Java classes and methods, although not all kind of story patterns can be translated automatically. Again, the kinds of analysis is limited in these approaches.

Kermeta [32] is an extension of EMOF (part of the MOF 2.0 specification) for specifying operational semantics. It enriches the EMOF metamodel with an action specification metamodel, introducing another new language to express specifications of algorithms. Simulation and execution possibilities are available for this approach.

Other works propose model transformation languages to specify the semantics of DSLs. For example, in [27], QVT is proposed to specify the semantics of OCL. Then, QVT rules are specified between models conforming to the same metamodel, representing in such a way the metamodel behavior, thus acquiring in-place transformations semantics (as graph transformations have). Analysis capabilities are not provided in this case.

Another interesting approach for defining semantic mappings in order to specify the semantics of a language is the semantic anchoring method developed at the Vanderbilt University [10]. Semantic anchoring relies on the use of well-defined “semantic units” of simple, well-understood constructs and on the use of model transformations that map higher level modeling constructs into configured semantic units. This approach uses Abstract State Machines as a semantic framework, and Microsoft’s Abstract State Machine Language (AsmL) and associated tools for programming, simulating and model checking ASM models [10]. One problem with this kind of approaches is that they normally force the introduction of yet another notation for specifying the mappings [33], or require the mappings to be proved to be correct. By using a single semantic domain (such as our proposal with Maude), we avoid the need of defining the so-called semantic mappings.

Other group of related works includes those that formalize the concepts of concrete metamodeling languages in Maude, such as UML or OCL. For example, RIVIERA [40] is a framework for the verification and simulation of UML class diagrams and statecharts. It is based on the representation of class and state models as terms in Maude modules that specify the UML metamodel; it makes a heavy use of the reflective capabilities of Maude, defining and handling most model operations at the metalevel. MOV A [42] is another Maude-based modeling framework for UML, which provides support for OCL constraint validation, OCL query evaluation, and OCL-based metrication. In MOV A, both UML class and object diagrams are formalized as MEL theories. MOMENT [5, 41] is a generic model management framework which uses Maude modules to automatically serialize software artifacts. It supports OCL queries and is also integrated in Eclipse. These systems use MEL and Maude to provide semantics to UML and MOF metamodeling frameworks.
Poernomo [34] also defines a formal metamodeling framework for MOF based on Constructive Type Theory, where models are defined as terms (token models), that can also be represented as types (type models) by means of a reflection mechanism. Similarly, Boronat and Meseguer have recently proposed an algebraic semantics of MOF in MEL [6], now also part of the MOMENT project [41], which provides a reflective, algebraic, and executable framework for metamodeling, with support for MOF and OCL. These two approaches currently provide support for the formal specification of the structural aspects of (MOF-based) models and metamodels, and for different kinds of model management operations (as we also support in our proposal, cf. [39]). However, they do not provide any specific approach to specify the dynamics of models, something specially important for achieving simulation and other kinds of behavioral analyses, as described in this paper.

7 Conclusions

According to the MDSD principles, models and metamodels become first-class citizens in the software engineering process. Several notations have been proposed to specify them, although the kind of formal and tool support they provide is always limited. In this paper we have shown how Maude provides an accurate way of specifying both the abstract syntax and the behavioral semantics of models and metamodels, and offers good tool support both for simulating and for reasoning about them.

Furthermore, Maude offers very good properties as a logical and semantic framework, in which many different logics and formalisms can be expressed [28]. Therefore, we expect that our proposal can be used to provide a rich semantic framework in which other proposals can be mapped, therefore allowing them to take advantage of the formal analysis methods and tools of Maude.

There are several lines of work in which we are currently engaged, or that we plan to address in the near future. For example, we are working on making Maude completely transparent to the user, i.e., we want users to benefit from Maude simulation and analysis capabilities, but working with the notation and tools they most like—i.e., we think that our proposal provides a perfect target semantic domain for providing behavioral semantics to modeling languages.

With this aim in mind, we are working on the integration of Maude with the Eclipse Modeling Framework (EMF) and graph grammars. Regarding the integration with EMF, we are defining ATL model transformations from EMF models conforming to arbitrary metamodels to the corresponding Maude specifications. These model transformations, together with a tool for representing rules that define the behavior of a metamodel in a visual fashion, will allow designers to simulate and analyze models and metamodels specified in the Eclipse platform.

Regarding graph grammars, we have already connected Maude with AToM3 [15] (a tool for multi-paradigm modeling based on graph grammars) to obtain the best of both platforms: allowing designers to work with domain specific concepts and their concrete syntax in a graphical way (with AToM3), while providing them with the advanced analysis tools of Maude [37].

Furthermore, we also want to study the expressiveness of our approach to add semantic information to model transformations (at the end of the day they are models, too), to be able to prove other kinds of properties, such as behavior preserving of model transformations [33, 35].

Finally, once we can model the basic behavior of models, we plan to explore the rich expressiveness of Maude to specify further behavioral aspects including, e.g., time, probabilities, quality of service and other non-functional properties, and combined hybrid behavioral semantics.

Acknowledgements The authors would like to thank the anonymous referees for their insightful comments and very constructive suggestions, that have helped us to significantly improve the contents and readability of the paper. This work has been partially supported by Spanish Research Projects TIN2008-03107, PET2006-0682-00 and P07-TIC-03184.

References


