

# Class Diagram Restructuring with GROOVE

Wietse Smid

University of Twente  
The Netherlands

w.b.t.smid@student.utwente.nl

Arend Rensink

University of Twente  
The Netherlands

arend.rensink@utwente.nl

This paper describes the GROOVE solution to the “Class Diagram Restructuring” case study of the Tool Transformation Contest 2013. We show that the visual rule formalism enables the required restructuring to be formulated in a very concise manner. Moreover, the GROOVE functionality for state space exploration allows checking confluence. Performance-wise, however, the solution does not scale well.

## 1 Introduction

This paper describes the GROOVE solution to the “Class Diagram Restructuring” case study of the Tool Transformation Contest 2013 [4].

GROOVE [5, 2] is designed to support state space exploration and model checking of dynamically evolving systems; however, as we will show in this contribution, it can also be used to demonstrate graph transformations as a basis for model transformation. GROOVE’s main usage in this solution (and indeed the core feature of GROOVE) is to formulate the required changes as transformation rules and explore all possible ways to apply those rules to a given graph, representing the initial class diagram.

**Graphs and rules.** In GROOVE, graphs are defined as nodes and directed edges with labels. Nodes have a special type of labels: the type label defines the type for a node, this is used in conjunction with a type graph system that preserves type safety. Furthermore, attributes are supported in the form of directed edges to value nodes.[3]

Rules enable the manipulation of graphs. Main rule features are:

**Readers:** Edges and nodes that must be present for a rule to be applicable, and are unaffected by the transformation. Readers are visually distinguished by black, continuous lines.

**Erasers:** Edges and nodes that (like readers) must be present for a rule to be applicable, but are deleted when the rule is applied. Erasers are visually distinguished by dashed blue lines.

**Creators:** Edges and nodes that are created when the rule is applied. Creators are visually distinguished by fat green lines.

**Embargoes:** Edges and nodes that are forbidden; i.e., when present they prevent a rule from being applicable. Embargoes are visually distinguished by fat, dotted red lines.

**Quantification:** Universal and Existential quantification are possible via “ $\forall$ ” and “ $\exists$ ” nodes. Other nodes can be bound to these quantifiers using edges labeled “@”. Quantifiers can also be nested using edges labeled “in”. For more details see [6].

In the Class Diagram Restructuring case study, the start graph was provided by an ECORE model. Since recently (see [1]), GROOVE can import and export ECORE models. When importing, the user can make choices regarding the representation of certain ECORE features that have no direct GROOVE counterpart. For instance, the models in this case study have opposing edges between Generalization and Entity (e.g.,

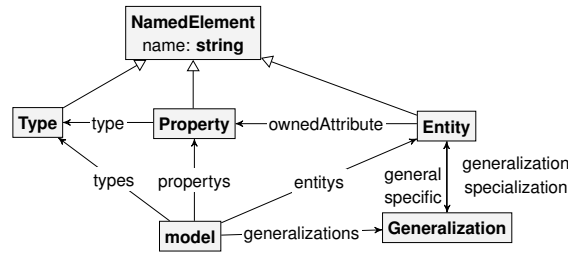


Figure 1: GROOVE type graph created from the Ecore metamodel

generalization vs. specific); in the chosen GROOVE representation, these are modelled by separate unidirectional edges. The actual type graph is shown in Fig. 1. Another possible choice would have been to introduce an intermediate node for each pair of concrete opposing edges; that blows up the model but retains more information, viz. the fact that the edges actually belong together.

To use GROOVE to actually carry out the transformations of this case study, the following steps are required:

1. Load the rule system into the Simulator (the GUI component of GROOVE);
2. Import the `.xmi` file of the input model using the Ecore importer;
3. Explore the state space, either manually (step by step) or automatically (completely);
4. Export the final state of the exploration using the Ecore exporter.

Alternatively, the exploration can be run from the command line using the Generator (the headless component of GROOVE).

## 2 Transformation Steps and Rule Applications

The case study calls for three distinct types of transformation steps; these should be executed to transform the given graph to its final state, viz. the restructured class diagram.

**Step 1: Pull Up common attributes of all direct subclasses** The first transformation step is by far the easiest and therefore has the simplest translation: see Fig. 2. Called `PULLUPDIRECT`, it searches for all `Property`s of which the containing `Entity` has a given `Entity` as its superclass. It then checks if all the `Entity`s with that superclass have the same `Property` in terms of type and name. The number of subclasses that have the `Property` must be larger than 1. The `Property`s are removed from the independent classes and a `Property` with the same type and name is created for the superclass. Note that the in-labelled (dotted) edge from the  $\exists$ -node to the  $\forall$ -node signifies nesting of the quantifiers: *for all* `Generalization+Entity`-pairs *there exists* a `Property`.

**Step 2: Create subclasses for duplicated attributes** To execute the creation of a subclass for duplicated attributes we divide the work into three consecutive rules. We guarantee the consecutive application of these sub steps by using embargoes in the rules on the temporary nodes that are created during previous steps.

First we apply the rule `SELECTSUBCLASS` (Fig. 3). Similar to `PULLUPDIRECT`, it matches `Entity`s that have a superclass and a `Property`. It then matches all the `Entity`s that fall under the superclass and have a `Property` with the same name and type. This explicitly does not mean that all the subclasses need

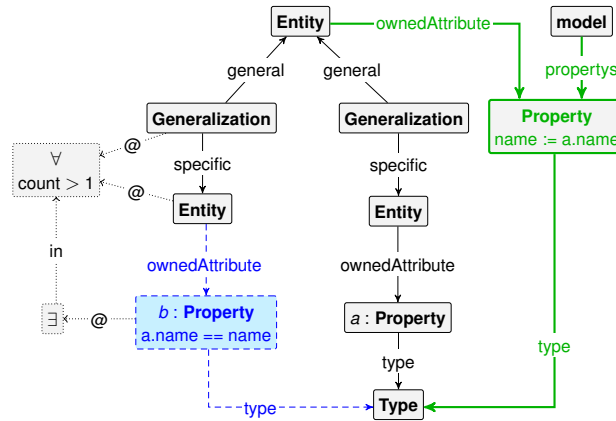


Figure 2: PULLUPDIRECT rule

this property. For this match it creates a `SelectSub` node for the selected superclass entity, and it links to all the names of the variables that are a candidate for upwards movement in the hierarchy.

Secondly we count for each of these candidates the number of occurrences, as we want to execute only the largest case. The rule `COUNTSUBCLASS` does exactly this. For each of the variables linked in the `SelectSub` node it takes the linked superclass `Entity` node and selects all the subclasses that have a variable of the name. For each of the linked names a `SelectedSubProperty` node is created that is linked to the properties, the type, and the name of the property. Furthermore the universal quantifier that is used to select the entities and properties for each of these linked names has a variable `count`, which records the number of matched items.

Finally we can then create the new class using a rule `CREATESUBCLASS` (omitted here for lack of space). It matches the `SelectedSubProperty` with the greatest number of elements. The generalizations that link the `Entity`s to the superclass are removed, as well as the properties themselves. A new `Entity`, subclass to the original superclass and superclass to the original subclasses is inserted, having a property of the same name and type as the matched property in `SelectedSubProperty`.

**Step 3: Create root class for duplicated attributes** Step 3 is actually a special case of step 2; only here we do not match on a common superclass, but on a common shared property, namely the absence of a superclass. The rules are therefore changed a bit, but the effective procedure matches that of step 2: first a root class is selected, then a count is performed for all suitable `Properties`, and finally the one with the maximum count is selected and a new root class is created for it.

**Extension to multiple inheritance** Because `GROOVE` is a general-purpose transformation tool, there is no specific reason in terms of graph transformations why multiple inheritance is impossible. There are of course a few things to consider.

Representing multiple inheritance is easy enough; the ecore type graph could allow for more than one general or specific to a generalization node. Trying to find if there is a shared superclass is then a case of finding a match where these edges exist.

Furthermore, a situation can be thought of where all subclasses of a class either have an identical property, or have a superclass with that property, with the added constraint that that superclass cannot

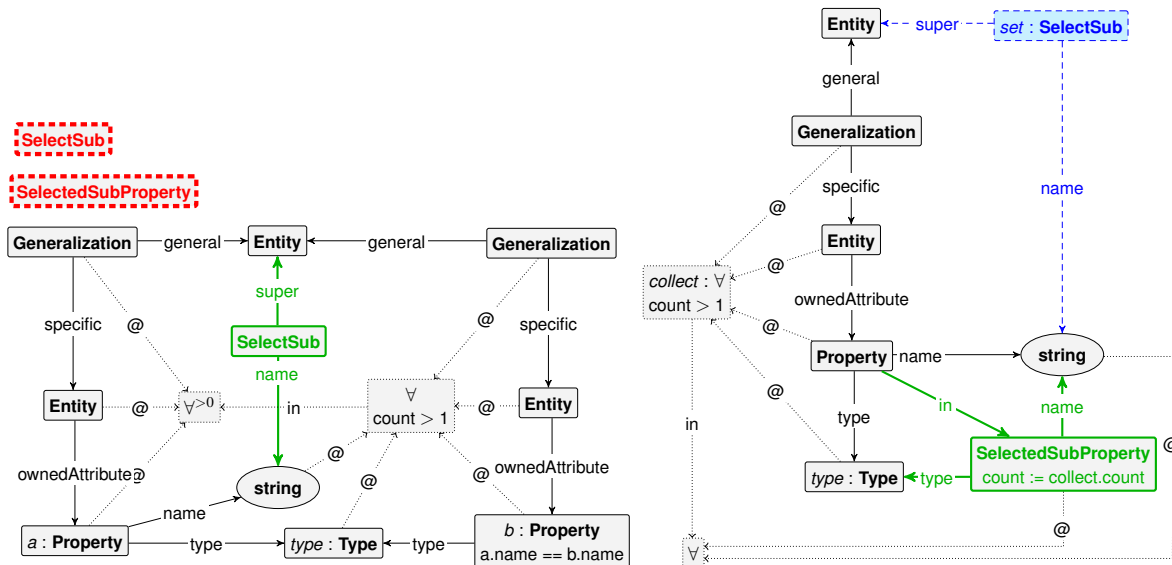


Figure 3: Rules for SELECTSUBCLASS and COUNTSUBCLASS

have subclasses outside of the original superclass. In this way, rules for multiple inheritance can get quite complicated; however, to create them is still eminently possible.

The visual interface that GROOVE provides is a useful tool for this job. All in all, constructing the rules for multiple inheritance is inherently difficult, but certainly possible with GROOVE.

### 3 Discussion and conclusion

GROOVE has its pros and cons as a tool for this case study. On the upside, GROOVE provides a few interesting features that were not required in the scope of the case description. The most important feature is the total state space exploration, this allows us to visualise things like confluence of a rule system as well as how a refactoring was reached.

Another positive point is the ease of use of the GROOVE tool. Because of its visual interface, designing rules and understanding them is relatively easy. Furthermore the rule application highlights the matched nodes and edges, allowing for direct feedback on created rules.

**Performance.** These advantages come at a price. Because GROOVE was created for a wider application area and for a different core functionality, even medium sized models pose a performance problem for GROOVE: the amount of resources needed to complete these examples outstrips what the Java VM can handle on a regular sized computer. Provided with 10G of memory on a 64-bit machine (Intel i7-2600 CPU @ 3.40GHz), the performance on the example cases is summarised in the following table:

Name	Size	#States	Time (s)
testcase1	14	3	< 1
testcase2	18	9	< 1
testcase3	5500	13	212
testcase2_1000	10000	Out of memory	

The size is the combined number of classes and attributes; the number of states is the set of consecutive graphs that were computed in the course of the exploration. Testcase 3 involves a very large combined

refactoring: 10 attributes are pulled up from 500 classes. Testcase 2\_1000 is a 1000-fold copy of testcase 2. Clearly, the performance of these larger cases is not practical.

**Rule scheduling.** The three step process for transformations 2 and 3 (see Sect. 2) has an inbuilt sequence that is enforced by the embargoes specifying the absence of the Selected\* nodes. This is a rather crude way of scheduling rule applications. GROOVE does provide another way to handle this, through control programs that explicitly specify that SELECTSUBCLASS, COUNTSUBCLASS and CREATESUBCLASS must always be performed in that order and never be interrupted by other rules. This feature has not been used here.

**Case ambiguities.** The Case Description leaves a few questions open. Firstly in the application of Rule 2 and Rule 3 it is posed that the first application of these rules should be on the property with the highest number of occurrences. There are however situations where there are two options of equal size which are not isomorphic, thus creating two final transformed graphs that are not confluent.

The other is the ambiguity in the naming of newly created Entitys. Although GROOVE has no real need for the names of Entitys — it can simply work with nameless entities because it only works on nodes and their edges — there is still the need for giving Entitys distinct and useful names. This is not a problem when exporting the graph from GROOVE to an xmi file, because GROOVE uses the internal IDs of the nodes for naming.

**Conclusion.** The provided solution can transform Class Diagrams according to the given transformation rules. Even though the solution lacks the performance to solve the largest test cases, there are other benefits not required from the case description. Confluence can be shown to hold in cases where the entire state space can be calculated. Furthermore the GROOVE tool has a high ease-of-use factor, and can be used for many more similar problems.

## References

- [1] Harold Bruintjes (2012): *Bridging GROOVE to the world using an abstracted language model*. Master's thesis, University of Twente. Available at <http://fmt.cs.utwente.nl/education/master/142/>.
- [2] Amir Hossein Ghamarian, Maarten J. de Mol, Arend Rensink, Eduardo Zambon & Maria V. Zimakova (2012): *Modelling and analysis using GROOVE*. *International journal on software tools for technology transfer* 14(1), pp. 15–40, doi:10.007/s10009-011-0186-x.
- [3] Harmen Kastenberg (2006): *Towards Attributed Graphs in Groove*. In: *Proceedings of the Workshop on Graph Transformation for Verification and Concurrency (GT-VC), Electronic Notes in Theoretical Computer Science* 154, Elsevier, pp. 47–54, doi:10.1016/j.entcs.2005.03.030.
- [4] Kevin Lano & Sefoukeh Kolahdouz-Rahimi (2013): *Case study: Class diagram restructuring*. In Pieter van Gorp, Christian Krause & Louis Rose, editors: *Sixth Transformation Tool Contest (TTC 2013)*, EPTCS. In this volume.
- [5] Arend Rensink (2004): *The GROOVE Simulator: A Tool for State Space Generation*. In J. L. Pfaltz, M. Nagl & B. Böhlen, editors: *Applications of Graph Transformations with Industrial Relevance (AGTIVE)*, LNCS 3062, Springer Verlag, pp. 479–485, doi:10.1007/978-3-540-25959-6\_40. Available at <http://groove.cs.utwente.nl/>.
- [6] Arend Rensink & Jan-Hendrik Kuperus (2009): *Repotting the geraniums: on nested graph transformation rules*. In A. Boronat & R. Heckel, editors: *Graph transformation and visual modelling techniques, York, U.K., Electronic Communications of the EASST 18*, EASST.