Reusable Idioms and Patterns in Graph Transformation Languages

Aditya Agrawal[1]   Attila Vizhanyo[2]   Zsolt Kalmar[3]   Feng Shi[4]
Anantha Narayanan[5]   Gabor Karsai[6]

Institute for Software Integrated Systems (ISIS)
Vanderbilt University
Nashville, TN 37235, USA

Abstract

Software engineering tools based on Graph Transformation techniques are becoming available, but their practical applicability is somewhat reduced by the lack of idioms and design patterns. Idioms and design patterns provide prototypical solutions for recurring design problems in software engineering, but their use can be easily extended into software development using graph transformation systems. In this paper we briefly present a simple graph transformation language: GREAT, and show how typical design problems that arise in the context of model transformations can be solved using its constructs. These solutions are similar to software design patterns, and intend to serve as the starting point for a more complete collection.

Key words: Graph Transformations, Design Patterns.

1   Introduction

The practical application of Graph Rewriting and Transformations (GRT) [4] is contingent upon the existence of mathematically well-founded, yet easy-to-use tools on one hand, and on the real-world engineering experience and knowledge about the use of the techniques on the other hand. With the arrival of the Model-Driven Architecture (MDA) [3], GRT is about to become

---

[1] `aditya@isis.vanderbilt.edu`
[2] `viza@isis.vanderbilt.edu`
[3] `kalmar@isis.vanderbilt.edu`
[4] `fengshi@isis.vanderbilt.edu`
[5] `ananth@isis.vanderbilt.edu`
[6] `gabor.karsai@vanderbilt.edu`

a technology that could be widely used in the industry. Although there have been a number of GRT tools developed [14,15,20,21], few tools have been used on practical development projects, and even less engineering experience has been accumulated and documented about the use of these tools.

We agree with the vision of MDA, where transformations on the artifacts produced during the design of software are an integral and essential part of the design process. We envision software development environments where model transformations are used to facilitate *design automation*. Transformations could take place at different phases of the process, for instance: (1) when design models are built and some activities (e.g. applying a design pattern) are best implemented by an automated tool that transforms the models [12],(2)when components are adapted to suit the needs of a particular design context [16], (3) when designs have to be transformed into models for analysis [10], (4) when code has to be generated from the models by instantiating code fragments [17]. As these transformations must be performed on design models (which are typed multi-graphs in the most general sense), GRT techniques are applicable.

There are (at least) two major motivations for using GRT in this context: (1) transformations could be complex, and a concise and precise language to program them is desirable, and (2) if the transformations are specified in a formal way (as the GRT technology allows it), we have the opportunity to reason about their properties, and the way they change the properties of the models they are applied to.

A language to write graph transformation programs in (and thus implement model transformations) should have a well-defined, yet simple syntax and semantics. However, there are common recurring tasks in model transformations that should not be directly supported by the language. Rather, they should be available as well-documented, reusable idioms and design patterns that solve recurring design problems. The difference between the two is that idioms are restricted to an application domain, while design patterns are domain-independent. In this paper, first we describe GReAT, a visual transformation language that supports explicitly sequenced graph transformation and rewriting operations. Next two domain-independent design patterns are described, followed by a description of a non-trivial idiom. The final sections discuss related and future work.

2   GReAT

The transformation language used to demonstrate the design patterns, algorithms and idioms is Graph Rewriting and Transformation language (GReAT). GReAT is a language that allows users to specify graph transformations in a graphical form with formal and executable semantics. In this paper only the

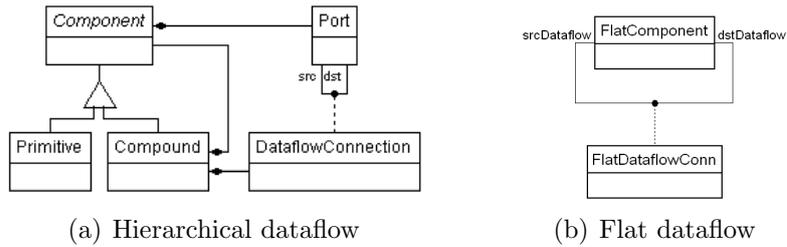(a) Hierarchical dataflow   (b) Flat dataflow

Fig. 1. Metamodels HDF and FDF

necessary language constructs are explained, [7] describes the full approach and support tools. The operational semantics of GReAT is formally defined in [9]. GReAT is based on the theoretical work of graph grammars and transformations [4,5,6] and belongs to the set of practical graph transformations systems, like AGG[15], PROGRES[14], FUJABA[20] and VIATRA[21]. GReAT can be divided into three parts:

  (i) Domain specification and heterogeneous transformations

 (ii) Graph transformation language

(iii) Control flow language.

2.1 Domain Specification and Heterogeneous Transformations

In graph transformation programs nodes and edges are typed, and these types form a type system that we call a domain. Many approaches have been introduced in the literature to capture graph domains. For instance, schemas are used in PROGRES while AGG uses type graphs. These are proprietary formats for the specification of the graph domain. We chose UML [1] class diagrams and the Object Constraint Language (OCL) [2] for such a specification because it is standardized and is at least as expressive as both schemas and type graphs. UML has been previously used for domain specification in FUJABA while MOF like models graphs have been used in VIATRA. Note that the UML class diagram plays the role of a graph grammar such that it can describe all the "legal" object networks that can be constructed within the domain.

From the UML class diagrams one can generate an object oriented API that can be used to implement the graphs, to traverse the input graph, and to construct the output graph. To support this, GReAT supports any number of domains in a transformation program.

For example, Figure 1(a) shows a UML class diagram that represents the domain of hierarchical data flow networks (HDF). A hierarchical data flow network consists of *Components*, *DataflowConnections* and *Ports*. A component can be either a *Compound* or a *Primitive* component, and a *Compound* component may contain other components. Components contain ports and
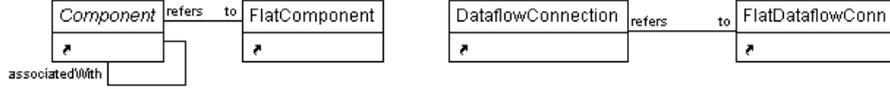
Fig. 2. A meta-model that introduces cross-links

directed dataflow associations between these ports represent the flow of data. Figure 1(b), represents a flat (non-hierarchical) dataflow language (FDF) with *FlatComponents* and dataflow connections between them. The hierarchical data flow network and flat representation will serves as an ongoing example throughout this paper.

In this paper, graph transformations are defined as transformations that manipulate graphs within the same domain, while graph rewriting is defined as a mapping between two different domains. A design challenge for GReAT was to provide a uniform syntax and semantics for both graph transformations and rewriting. This problem is tackled in GReAT by allowing the user to compose source and target metamodels by defining temporary vertex and edge types that can span across multiple domains and will be used temporarily during the transformation. For example, Figure 2 shows a metamodel that defines associations/edges between HDF and FDF. *Component* and *DataflowConnection* are classes from Figure 1(a) while the *FlatComponent* and *FlatDataflow* are classes from Figure 1(b). This metamodel defines three types of edges. There is a *refersTo* edge type that can exist between *Component* and *FlatComponent*, and between *DataflowConnection* and *FlatDataflowConn*. Another edge type *associatedWith* is defined and it can link *Component* objects. By composing the domains using temporary cross-links we are able to tie the different domains together to make a larger, heterogeneous domain that encompasses all the domains and cross-references. This approach is similar to reference nodes and edges in VIATRA[21].

2.2 Graph Transformation Language

The heart of GReAT is the graph transformation language. It was inspired by many previous efforts such as [4,5,6]. It defines the basic transformation entity: a production/rule. A production contains a pattern graph, in which each pattern object: a vertex (or an edge) conforms to a type: a class (or an association) from the metamodel. Apart from this, each pattern object has another attribute that specifies the role it plays in the transformation. There are three roles that a pattern object can play:

(i) *bind*: The object is used only to match objects in the graph.
(ii) *delete*: The object is used to match objects, but once the match is computed, the objects are deleted.
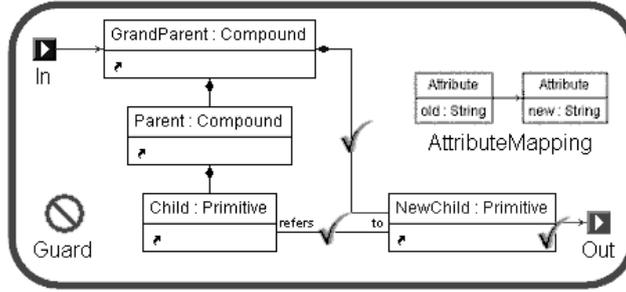(iii) *new*: New objects are created after the match is computed.

Fig. 3. Example rule with patterns, guards and attribute mapping

The execution of a rule involves matching every pattern object marked either *bind* or *delete*. The pattern matcher will return all possible matches for the given pattern graph and host graph. Then for each match the pattern objects marked *delete* are removed and then the objects marked *new* are created.

Pre-conditions are often required for additional constraints on the transformation application. In GReAT, OCL constraints can be used to specify guards (see Figure 3). These guards are evaluated on the matches before the actions are applied. There is also a need to provide values to attributes of newly created objects and/or modify attributes of existing objects. "Attribute mapping" (AM) is a specification of such attribute manipulation and is executed after the transformation is applied. The AM specification consists of code based on a C++ API that accesses the graph objects. Figure 3 shows an example rule. Each object in the pattern graph refers to a class in the heterogeneous metamodel. The semantic meaning of this reference is that the pattern object should be matched to a graph object that is an instance of the class represented by the metamodel entity. The default action of the pattern objects is bind. The new action is denoted by a tick mark on the pattern vertex (see the vertex *NewChild* in Figure 3). Delete is represented using a cross mark (not shown in Figure). Our action specification is similar to erasers and creators in Groove [23], or to diagram elements marked destroy and create in FUJABA [20]. The In and Out icons in the Figure are used for passing graph objects between rules and will be discussed in detail in the next section.

2.3 Controlled Graph Rewriting and Transformation

GReAT has a high-level control flow language built on top of the graph transformation language with the following constructs for improving the efficiency of the transformations: (1) pivoting and (2) sequencing. In this paper these issues will be briefly touched upon, for a complete discussion please refer to [8].

The performance of the pattern matching can be significantly increased if some of the pattern variables are bound to elements of the host graph

before the matching algorithm is started (effectively providing a context for the search). The initial matches, called pivots, are provided to a transformation rule with the help of ports that form the input and output interface for each transformation step. Thus a transformation rule is similar to a function, which is applied to the set of bindings received through the input ports and results in a set of bindings over the output ports. For a transformation to be executed graph objects must be supplied to each port in the input interface. In Figure 3 the In and Out icons are input and output ports, respectively. Input ports provide the initial match to the pattern matcher while output ports are used to extract graph objects from the rule so that they can be passed along to the next rule. The rules thus operate on packets, which are defined as sets of (port, host graph object) pairs. One execution of rule consumes all input packets incident on it to produce a number of output packets.

Explicit sequencing of rules and a high-level control flow language allows the precise control of transformations and thus helps manage the complexity of the transformation and allows users to write efficient transformations. The control flow language supports the following features:

(i) Sequencing: Rules can be sequenced to fire one after another.

(ii) Non-Determinism: The execution order of "parallel" rules on a sequential machine is non-deterministic.

(iii) Hierarchy: Compound rules can contain other rules.

(iv) Rule reuse: The same rule can be called from different parts of the transformation specification.

(v) Recursion: A high-level rule can (directly or indirectly) call itself.

(vi) Test/Case: A branching construct to choose between control flow paths.

Sequencing is used to specify an order of execution for a set of transformation rules. For example, Figure 5 shows a sequence of rules, "HasComponents" and "Call: CollectPrimitives" are executed sequentially which is in parallel with the rule "IsPrimitive". Story diagrams [22] in FUJABA have a similar sequencing construct. Hierarchy is also shown in Figure 5, where the above mentioned rules are all contained in a compound rule called the "CollectPrimitives".

A test/case construct is used to choose between different execution paths, similar to the 'if' statement in programming languages. Figure 9, contains a test called "TestProxyExistence" that contains two cases (shown in Figure 10). The test will first try case "HasProxy", if "HasProxy" succeeds then the outputs will be passed to the respective output ports and similarly for "NoProxy". Once all inputs have been evaluated the next rules in the sequence will be executed. Due to space constraints only a brief explanation is provided in this section. For a detailed discussion of these constructs please refer to [7]

and for the formal semantics of the language please refer to [9].

3 Patterns and Idioms for Reusable Graph Transformations

Software design is commonly regarded as the most difficult stage in the software development cycle. However, some design challenges are common and have been faced by many software designers. Over the years elegant solutions to such problems have been identified and implemented. When such a design solution is formally documented, identifying the participating elements, their roles and collaborations, and the distribution of responsibilities, a design pattern arises. A design pattern allows some aspect of system structure vary independently of other aspects, thereby promoting robustness and domain-independent reuse [11].

The same driving forces exist in the area of programming graph transformations and in the following section a few transformation patterns, algorithms and idioms will be introduced. Each pattern/idiom will be described in a uniform way with the following structure: (1) Motivation: a problem, where the need for the pattern arises. (2) Applicability: the general class of problems where the pattern is applicable. (3) Structure: the abstract specification of the pattern. (4) Benefits: the advantages of applying the pattern. (5) Known uses: a set of transformations where the pattern has been known to be applied. (6) Limitations: a set of limitations to the applicability of the patterns.

Though these design patterns are applicable to a broad variety of transformations a single motivating example is used in this paper and will be carried throughout the design pattern discussion. The example is the flattening of hierarchical dataflow (HDF) to a flat dataflow (FDF) representation. In Figure 1(a) the meta-model of the HDF has been presented with primitives that capture dataflow behavior, and compounds that are only used for encapsulating other components. The aim is to convert the tree structure of HDF to an FDF representation (see Figure 1(b)) while preserving the dataflow connectivity. A simple algorithm is the following: (1) collect all primitive nodes and copy them to FDF, (2) trace the dataflow connection from each port in each primitive to a corresponding target primitive port, (3) replace this trace by a single dataflow association in FDF.

3.1 The Leaf Collector Pattern

Motivation: In step 1 of the HDF flattening algorithm, a requirement is to collect all leaf nodes in the hierarchy. For example, in Figure 4, given the root component C we need to find all leaf primitives P_{121}, P_{11}, P_{21}, P_{22}. Figure 5 shows the rules that collect all the primitives in a given HDF hierarchy.
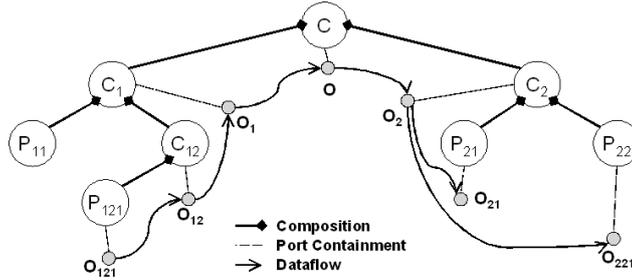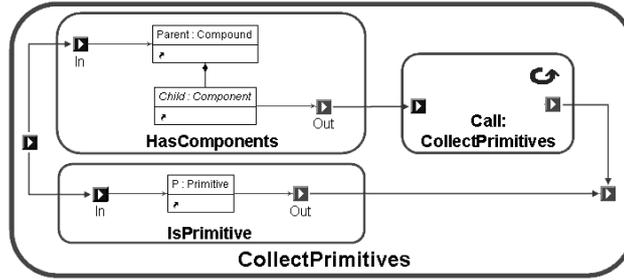
7

Fig. 4. An HDF model

Fig. 5. Collecting primitives in HDF

The top level rule "CollectPrimitives", gets as an input the root object of the hierarchy. It calls "HasComponents" that collects all direct children and on each child a recursive call to "CollectPrimitives" is made. If the input to "CollectPrimitives" is a primitive, then the "IsPrimitive" rule will succeed and its input will be passed to the output. At the end of the recursion, all primitives will be available at the output of the top level call.

Applicability: From a starting object, traversal of a particular kind of directed association is required till leaf objects are reached. Leaf objects are defined as objects, from which the association cannot be traversed any more.

Structure: The participants of this pattern are shown in Figure 6. The "GetDirectNeighbors" rule is responsible for collecting all the direct neighbors of the input object. The "IsLeaf" rule is responsible for identifying if the input object is a leaf. This is achieved by a pattern with pattern cardinality equal to zero (see arrow in "IsLeaf"). A cardinality of zero on a pattern edge means that no such edge should exist between the two objects. In other transformation languages this is referred as a negative pattern edge. This implementation of "IsLeaf" is more general than the one seen in the dataflow example.

Benefits: The traversal scheme and the leaf recognition are independent of each other. The processing of the leaves is separate from collecting the leaves and is typically done in a rule following the leaf collector pattern, and thus leaf collection and leaf processing can vary independently.

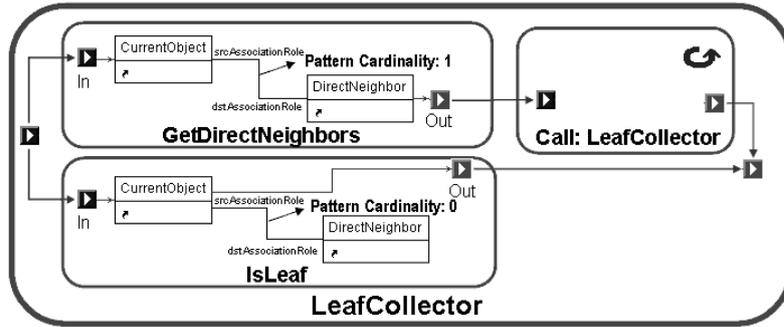*Known Uses:* The leaf collector pattern has been used in (1) hierarchical

Fig. 6. The structure of the Leaf Collector pattern

statemachine to finite statemachine transformation to get all leaf states, (2) Matlab's Simulink/Stateflow to hybrid automata transformation. Starting from a Simulink port to find all ports at the end of the dataflow connection chain. For example, in Figure 4, given port O_{121}, it would find O_{21} and O_{221}. (3) Embedded Systems Modeling Language (ESML) [18] to Task Network Architecture (TNA) transformation to collect all tasks in ESML where the tasks are hierarchical.

Limitations: The leaf collector pattern can only be applied graphs that do not have cycles in them.

3.2 Transitive Closure

Motivation: In the FDF a data dependency analysis needs to be performed. For such an analysis the transitive closure [19] of the dataflow connection needs to be found. Figure 7 shows the transformation to compute the transitive closure on FDF. The first step of the algorithm is to find all neighbors of those FDF components that do not have any incoming associations (source). This is achieved in "FindSource" using the zero cardinality association pattern. Then the next neighbors are used as the initial value for the "TransitiveClosureStep". In one "TransitiveClosureStep", given a set of components, all their next and previous neighbors are found. If there is no association from the previous neighbor to the next neighbor, a new association is created between them. The "TransitiveClosureStep" is called again with the set of next neighbors as input.

Applicability: In all situations where the transitive closure needs to be computed.

Structure: The algorithm presented in Figure 7 works only on directed acyclic graphs. Preprocessing steps can be used to convert an arbitrary graph into a directed acyclic graph by reducing all strongly connected components to a single vertex. In the general case the type of the association and the type of the vertex can be changed to suit the requirements of the problem.
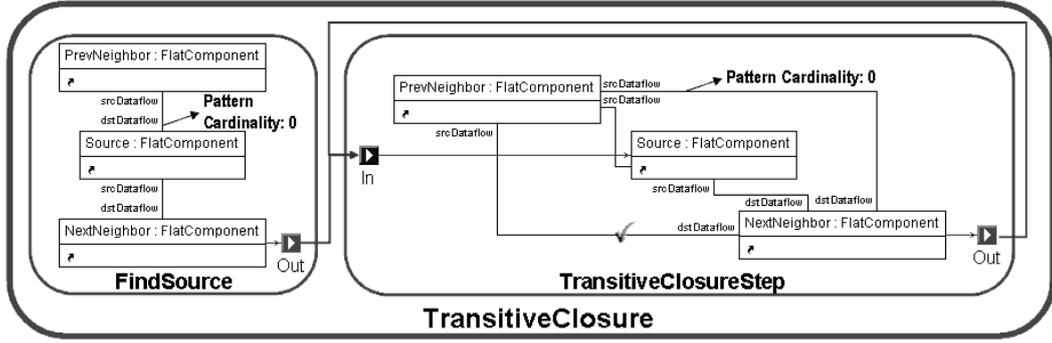
9

Fig. 7. Transformation rules to find the transitive closure
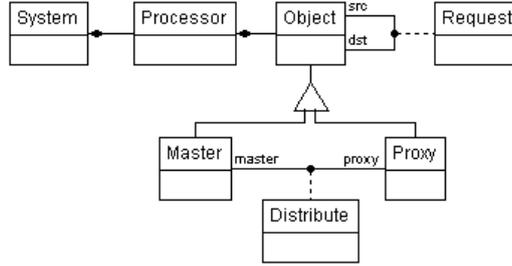

Fig. 8. UML class diagram of a general distributed system

Known Uses: Has been used to perform reachability analysis on Embedded Systems Modeling Language (ESML) [18] and other distributed and parallel systems. Transitive closure can also be used in the construction of a parsing automata.

Limitations: The transitive closure pattern can only be applied to directed acyclic graphs.

3.3 The Proxy Generator Idiom

So far this paper has discussed design patterns that have applicability to a large set of problem domains. This section will discuss the proxy generator design idiom, which is also reusable, but is restricted to a particular problem domain.

Motivation: A model for a distributed service based system (DSBS) is shown in Figure 8. The distributed system (*System*) is composed of multiple processors (*Processor*), each processor hosts different objects (*Object*), and objects may request service (*Request*) from other objects. An object could be a proxy (*Proxy*) of a remote master object (*Master*) on a local processor, and such a relationship could be represented by the association *Distribute* between two objects. Therefore a proxy is a placeholder for another object and provides access control to it. In such distributed systems, in order to reduce network traffic as well as to abstract network communication from
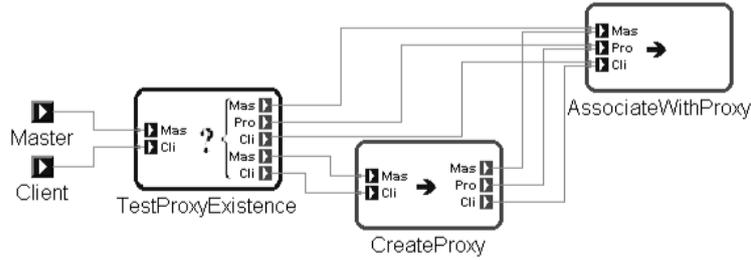
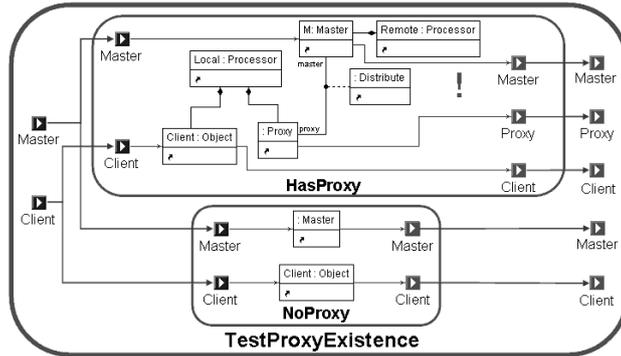Fig. 9. Transformation rules for the proxy generator idiom

Fig. 10. TestProxyExistence

object interactions, we can use proxies on each processor to represent remote components locally. This optimization can be performed by identifying the need for proxies using static analysis methods on the object network, and then automatically generating them.

Figure 9 shows the transformation that identifies the need for proxies and creates them. For each *Master*, *Client* pair, such that the *Client* needs to make requests on the remote *Master*, the first step is to determine whether the *Master* has a proxy on the client's processor. This is done in "TestProxyExistence". If it succeeds in finding a proxy, then "AssociateWithProxy" is called, which replaces the request to the master with a request to the local proxy. Otherwise, "CreateProxy" is called, which creates the equivalent proxy on the client's processor, followed by a call to "AssociateWithProxy".

Step 1. Determine proxy presence on local processor: This is achieved using a test called "TestProxyExistence" depicted in Figure 10. It has two cases, "HasProxy" which checks the existence of a proxy for the *Master* on the client's processor. If "HasProxy" is successful, the *Master*, *Proxy* and *Client* are passed to the "AssociateWithProxy" rule. Otherwise the "NoProxy" case is called, that always succeeds and will pass the *Master*, *Client* pair to the next rule.

Step 2. Create the proxy (optional): In the rule "CreateProxy" as shown in Figure 11(a), given the *Master* and *Client*, a *Proxy* is created on the

11

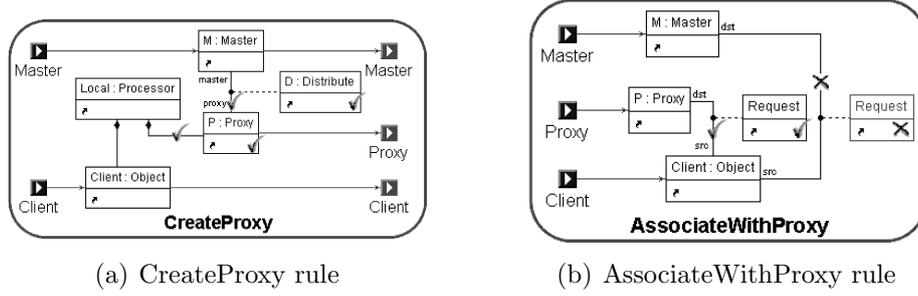(a) CreateProxy rule   (b) AssociateWithProxy rule

Fig. 11. Create and associate proxy

client's local processor along with a *Distribute* association with the *Master*. The newly created *Proxy*, its corresponding *Master* and the *Client* are passed to the "AssociateWithProxy" rule.

Step 3. Migrate the services: In the rule "AssociateWithProxy" as shown in Figure 11(b), given the *Master* and the corresponding *Proxy*, the *Client*'s request to the *Master* is replaced by the *Client*'s request to the *Proxy*.

Applicability: In any distributed system where remote interactions need to be abstracted and optimized. These interactions could be service/request, event source/sink, and dataflow interactions. The idiom can be used to perform a static analysis on a network of distributed objects for optimization, or can be used at runtime when the need for remote interaction arises.

Structure: In the general case the idiom has three distinct tasks. (1) Given a master-client pair, test for the presence of the proxy. (2) If a proxy is present, remove the client-master association and create a client-proxy association. (3) If a proxy is not present, first create a proxy and link it with the master, then remove the client-master association and create a client-proxy association.

Consequences: The pattern separates the proxy identification, proxy creation and proxy association steps such that each can be modified independently.

Known Uses: Used in the Embedded System Modeling Language (ESML) [18].

Limitations: The proxy generator idiom expects a client and master pair such that a request exists between the two.

4 Related Work

Software design patterns have been documented since the early 1990's. [11] is considered the ground-breaking work on design patterns.

Graph transformations have been proposed to automate the application of software design patterns. [12] describes an approach which uses graph queries and graph rewriting rules to apply software design patterns. In [13]

the suitability of Graph Rewriting Systems (GRS) for the specification and execution of complex transformations on the static aspects of design is argued. Examples in the paper illustrate how GRS can be used to implement the design transformations proposed by MDA. These approaches try to automate the application of software design patterns. Whereas this paper tries to discover transformation design patterns that are solutions to common transformation problems and can be used while designing graph transformations.

5 Conclusions

We have shown how the idea of design patterns applies in the context of graph transformation programs. We demonstrated the concept using three (domain-independent) design patterns and one (domain-dependent) idiom. We believe that documenting design patterns in this manner is useful for practitioners of graph transformations and we call the community to contribute to this body of knowledge. Further research in this area could lead to comprehensive compilation of design patterns and idioms that could be used to educate developers using the graph transformation technology.

6 Acknowledgement

References

[1] J. Rumbaugh, I. Jacobson, and G. Booch, "The Unified Modeling Language Reference Manual", Addison-Wesley, 1998.

[2] Object Management Group, Object Constraint Language Specification, OMG Document formal/01-9-77. September 2001.

[3] "The Model-Driven Architecture", http://www.omg.org/mda/, OMG, Needham, MA, 2002.

[4] Grzegorz Rozenberg, "Handbook of Graph Grammars and Computing by Graph Transformation", World Scientific Publishing Co. Pte. Ltd., 1997.

[5] Blostein D., Schürr A., "Computing with Graphs and Graph Rewriting", Technical Report AIB 97-8, Fachgruppe Informatik, RWTH Aachen, Germany.

[6] H. Gottler, "Attributed graph grammars for graphics", H. Ehrig, M. Nagl, and G. Rosenberg, editors, Graph Grammars and their Application lo Computer Science, LNCS 153, pages 130-142, Springer-Verlag, 1982.

[7] Agrawal A., Karsai G., Shi F.: "Graph Transformations on Domain-Specific Models", Technical report ISIS-03-403, Vanderbilt University, November, 2003.

[8] Vizhanyo A., Agrawal A., Shi F.: "Towards Generation of High-performance Transformations", Generative Programming and Component Engineering, (submitted), Vancouver, Canada, October 24, 2004.

[9] Karsai G., Agrawal A., Shi F., Sprinkle J.: "On the Use of Graph Transformations for the Formal Specification of Model Interpreters", Journal of Universal Computer Science, Special issue on Formal Specification of CBS, 2003.

[10] Agrawal A., Simon G., Karsai G.: "Semantic Translation of Simulink/Stateflow models to Hybrid Automata using Graph Transformations", International Workshop on Graph Transformation and Visual Modeling Techniques, Barcelona, Spain, March 27, 2004.

[11] Erich Gamma, Richard Helm, Ralph Johnson, and John Vlissides: Design Patterns: Elements of Reusable Object-Oriented Software Addison-Wesley, 1995

[12] Angsgar Radermacher. Support for Design Patterns Through Graph Transformation Tools. AGTIVE 1999: pp. 111-126

[13] Alexander Christoph. Graph Rewrite Systems for Software Design Transformations; Objects, Components, Architectures, Services, and Applications for a Networked World: International Conference NetObjectDays, NODe 2002, Erfurt, Germany, October 7-10, 2002. pp. 76 - 86

[14] A. Schürr. PROGRES: A VHL-language based on graph grammars. In in Proc. 4th Int. Workshop on Graph-Grammars and Their Application to Computer Science, number 532 in LNCS, pages 641–659. Springer-Verlag, 1991. T

[15] Taentzer, G.: AGG: A Tool Enviroment for Algebraic Graph Transformation, in Proc. of Applications of Graph Transformation with Industrial Relevance, Kerkrade, The Netherlands, LNCS, Springer, 2000.

[16] U.Assmann: "Invasive Software Composition", Springer, 2003.

[17] R. Alur, F. Ivancic, J. Kim, I. Lee and O. Sokolsky: "Generating Embedded Software from Hierarchical Hybrid Models", in Proceeding of ACM SIGPLAN Languages, Compilers, and Tools for Embedded Systems (LCTES'03), San Diego, California, June 11-13, 2003.

[18] G. Karsai, S. Neema, B. Abbott, D. Sharp. "A Modeling Language and its Supporting Tools for Avionics Systems", 21st Digital Avionics Systems Conference, August 2002.

[19] Kurt Mehlhorn, "Graph algorithms and NP-completeness", Springer-Verlag New York, Inc., New York, NY, USA, 1984.

[20] U. Nickel and J. Niere and A. Zündorf, "Tool demonstration: The FUJABA environment", Proc. ICSE: The 22nd International Conference on Software Engineering, Limerick, Ireland, ACM Press, 200.

[21] D. Varró and G. Varró and A. Pataricza, Designing the Automatic Transformation of Visual Languages, volume 44, Elsevier, pages 205–227, Science of Computer Programming, 2002.

[22] T. Fischer and J. Niere and L. Torunski and A.Zündorf, "Story Diagrams: A new Graph Transformation Language based on UML and Java, Proc. Theory and Application to Graph Transformations , TAGT'98.

[23] A. Rensink, The GROOVE Simulator: A tool for state space generation, AGTIVE, Lecture Notes in Computer Science, Springer, 2003.