Development and Evaluation of a Metamodel to Define Modeling Syntaxes for CouchEdit

Florian Lappe

Universität Ulm

Fakultät für
Ingenieurwissenschaften
und Informatik

Institut für
Programmiermethodik
und Compilerbau

September 2020

Bachelorarbeit im
Studiengang Informatik

Abstract

Modeling has become an important factor in many of today's development workflows. Especially recent developments in the area of Model Driven Engineering (MDE) as well as Business Process Modeling, have increased the need for syntactically correct models. Therefore graphical modeling editors with syntax parsing capabilities are integral. But most modern model editors provide sub par user experience, resulting from tight coupling between graphical and abstract notations.

COUCHEDIT is a novel modeling language agnostic framework that strives to improve editor usability. To this end it introduces an architecture that decouples graphical and abstract representations. However, in its current implementation, configuring COUCHEDIT for different modeling languages is a laborious and error prone task.

This thesis proposes an architecture that can be used to define modeling language configurations for COUCHEDIT. The aim of this architecture is to increase developer accessibility by reducing the amount of code as well as framework knowledge needed to define new configurations.

Gutachter: Prof. Dr. Matthias Tichy
Betreuer: Dr. Alexander Raschke

Fassung October 5, 2020

Satz: PDF-LaTeX 2_ε

Contents

1 Introduction

Modeling languages have long played an important role in software engineering. Well designed models can abstract complex systems and provide visual aid in understanding them. Furthermore in form of the Business Process Model Notation (BPMN) they are used to define and automate processes. Today, with research in the area of Model Driven Engineering (MDE), a paradigm centered around models, with the intent to generate code bases and whole systems from them, the importance of models is rising even more.

As these tasks require syntactic correctness of used models, modeling tools become an essential part of an engineer's workflow. Especially visual modeling tools provide in theory, an intuitive and user friendly way to design models. But current graphical modeling tools tend to constraint users in unintuitive ways and deliver sub par user experience (UX). this usually arises from a tight coupling between a modeling tools user interface (UI) and the underlying model. As the model's syntax is usually inflexible, the UI has to make restrictions to adhere to this syntax. this often creates problems for the user, for example connections can only be drawn between two existing states, or deleting a node will result in all its children being deleted as well.

To amend these usability woes, L. Nachreiner proposed a novel modeling framework, called CouchEdit [1]. This framework decouples user interface and model syntax by introducing different models for both. Instead of relying directly on the syntax of the model that is being designed, in the CouchEdit architecture the user interface is using a render model that specifies only basic graphic objects. On the other hand, the actual abstract model syntax now stands on its own. To connect these two models, a third metamodel is Utilized. CouchEdit at its core was designed to be general purpose, meaning it can be rewritten to adhere to any model syntax. But to realize this in the current implementation, the source code has to be changed directly, which is error prone, convoluted and requires understanding of CouchEdit's internal architecture.

To create a more developer friendly and flexible way of adapting CouchEdit to different modeling syntaxes, this design research proposes a new metamodel, that can be used to create modeling syntax definitions which are usable by CouchEdit. Furthermore a conceptual parser is presented, that provides proof of concept on how this newly developed metamodel interacts with the CouchEdit architecture.

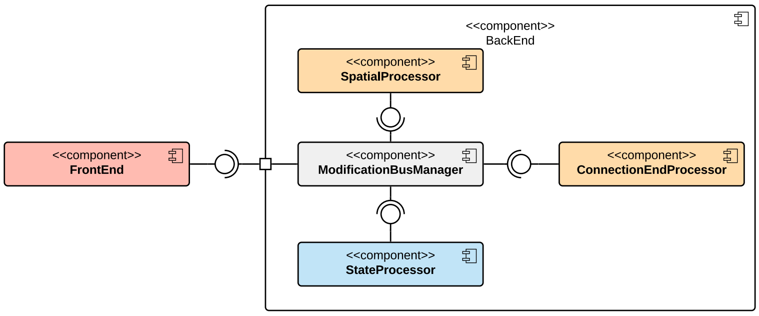


Figure 1.1: Publish Subscribe architecture of CouchEdit

1.1 Problem Statement

A general purpose framework should be configurable for multiple use cases in its designated domain. CouchEdit as a general purpose graphical modeling framework thus should be configurable for multiple modeling syntaxes. Technically this is possible, as long as one has access to the source code. But this would mean, every time CouchEdit has to support a new modeling syntax, manual changes in the source code have to be made and the project has to be compiled from sources. The implementation of a configuration parser, that can interpret modelling syntax definitions at runtime would thus increase flexibility. Furthermore a well designed metamodel could reduce the amount of knowledge that is needed about the CouchEdit framework, when configuring new model syntaxes.

As a relaxed conformance editing framework, CouchEdit poses special architectural requirements. It has to allow for temporary inconsistencies between concrete syntax (what the user draws) and abstract syntax (what the underlying Model actually looks like). As the concrete representation does not always have to map to a syntactic correct abstract model, the user has more freedom during the modeling process (e.g. dangling transitions).

CouchEdit achieves this by building upon the architecture concept of clear separation between concrete and abstract syntax, proposed by Y. Van Tendeloo et al. [2]. Internally, CouchEdit builds a hypergraph, that maps the given concrete syntax to all possible abstract syntaxes (interpretation of a concrete syntax can be ambiguous and thus multiple abstract syntaxes can be possible). To build this Hypergraph, a set of processors is employed, which are connected in a reactive publish and subscribe pattern. All processors (and the user interface) are subscribed to a bus (Figure 1.1). If a change (diff) is published to the bus (e.g. the user adds a node to the concrete syntax), all processors that are subscribed to this type of diff are notified and calculate new resulting diffs, these new diffs are then also published to the bus and all processors interested in them are invoked as well.

Some of these processors are needed for every type of syntax model, for example the spatial processor, that processes how nodes in the concrete syntax are positioned to each other (above, besides, etc.). Other processors are specific to the given modeling syntax, for example a Statechart syntax would require a state processor, that processes if a given graphical object represents a state (usually true if the given node is a rectangle with rounded edges).

A modeling syntax parser for CouchEdit would have to generate these syntax specific processors, while considering multiple design constraints that result from this architecture. Thus a metamodel is needed that can define the desired abstract modeling syntax and specify how a concrete representation can be mapped to this abstract syntax. While there is ongoing research in the area of relaxed conformance editing and how to link concrete and abstract syntax, it does not immediately become clear what such a metamodel can look like.

1.2 Purpose of this Study

The primary purpose of this study was to develop a metamodel for the CouchEdit framework. This metamodel is supposed to provide a comprehensive and easy to use way for defining new modeling languages. To evaluate the applicability of this metamodel, a prototypical code generator was implemented, that can comprehend this newly designed metamodel and translate it into a source code implementation. Furthermore a simple domain specific language was developed, that eases the process of specifying instances of this metamodel

This extension of the CouchEdit framework is supposed to improve developer accessibility and framework flexibility. While a code generator means that the system still has to be recompiled for every modeling syntax, it still becomes easier to support new modeling syntaxes as the metamodel abstracts away from the actual source code implementation and thus requires less knowledge about the CouchEdit framework, as well as fewer lines of code.

1.3 Research Questions

The primary goal of this research was the development of a metamodel for the CouchEdit framework. To this end, the following questions were defined as a means to guide the research.

RQ1: Is it possible to define a concise metamodel that covers CouchEdit's feature set?

If this is not possible, the following question has to be answered.

RQ1.1: How could the feature set be narrowed down, without impacting the frameworks capabilities too much?

RQ2: What requirements does the CouchEdit architecture impose onto a metamodel?

The architectures characteristics and quirks have to be considered when designing the language. For example, the publish and subscribe pattern, could make it easy to introduce nonterminating processing chains, which the metamodel should prevent where possible.

RQ3: How applicable is the designed metamodel for defining new modeling syntaxes?

RQ3.1 Is it possible to express common modeling syntax concepts in a clear and concise way?

RQ3.2 Are complex concepts still expressible without causing to much convolution?

1.4 Thesis Structure

This thesis is organized in the following structure: Chapter 2, lays out fundamentals required in the rest of the thesis. This includes a comprehensive summary of the COUCHEDIT architecture as well as modeling theory. There after Chapter 3 summarizes related work in the relevant fields. Following that, in Chapter 4 the developed metamodel is described in detail. The prototypical implementation that was developed alongside the metamodel, is then laid out in Chapter 5. Chapter 6 provides considerations about the evaluation of the artifact, as well as the evaluation results. Finally Chapter 7 gives perspectives on the future of this project, as well as takeaways from this research.

2 Fundamentals

this chapter lays out the knowledge foundation, required in later sections. First an outline of the applied research process is given. following that, a comprehensive view of the COUCHEDIT architecture is given. This will be needed later on to understand certain design decisions. Thereafter, a foundation of modeling languages is laid out, as well as some modeling syntaxes. Finally, a summary of related work and their impact on this thesis will be presented.

2.1 Methodology

As this research strived to develop a new metamodel suitable for the COUCHEDIT framework, it was conducted in accordance to the Design Science Research (DSR) approach. According to V. Vaishnavi et al. a design science research process consists of five steps [3].

2.1.1 Problem Awareness

The first step of a design science research is the identification of existing problems. As specified in Section 1.1, it was identified that the current COUCHEDIT implementation lacks a developer friendly way to configure it for different modeling syntaxes and that it is unclear how a metamodel for this architecture would look like.

2.1.2 Suggestion

With a clear definition of the problem, objectives can be proposed which have to be achieved in order to solve this problem. The first objective of this research is to develop a metamodel for the COUCHEDIT architecture. To be more precise, a metamodel is to be designed, that can be used to specify model syntax definitions which map concrete graphical syntaxes to Abstract syntax models and is applicable to COUCHEDIT's architecture. The second objective is to implement a prototype that provides proof of concept for the applicability of the design metamodel.

2.1.3 Development

The primary goal of a design science research is the development of artifacts. The first artifact to be developed in this research will be a metamodel, that can be used to define new modeling syntaxes for a relaxed conformance editor and is applicable to COUCHEDIT's architecture. The second artifact that is to be developed, is a prototypical code generator that can translate the developed metamodel into a COUCHEDIT implementation which will be able to process the defined modeling syntax. Furthermore it was decided to implement a simple DSL that eases the process of defining an instance of the given metamodel.

To this end, the first sub step of the development stage is a comprehensive analysis of COUCHEDIT's architecture. In his work L. Nachreiner describes in detail, which modeling features the framework covers and how they are implemented [1]. The goal is to develop an approach that can streamline the process of defining modeling syntaxes while adhering to COUCHEDIT's specification. Related work defined in Chapter 3 serves as a basis from which an initial suitable metamodel can be derived.

in the next sub step, the code generator will be implemented, this is an iterative step. first, an implementation will be developed on the basis of the designed metamodel, this should reveal further requirements imposed by COUCHEDIT's architecture (RQ2). The metamodel then has to be adjusted to account for these requirements, which in turn requires changes of the implementation until no further requirements can be deduced. Because of time constraints it is not possible to implement all of the metamodel's features, instead the prototype covers a minimal feature set that still suffices to demonstrate the metamodel's applicability for selected modeling syntaxes.

2.1.4 Evaluation

Now that the desired artifacts are fully developed it has to be evaluated how well they do their designated task. The metamodel as the primary artifact of this research has to be evaluated in terms of its applicability to its designated domain. To this end, the metamodel will be demonstrated on the basis of different modeling syntaxes, this should highlight areas in which the metamodel's design excels, as well as design flaws and limitations (RQ3). If time allows, potential flaws can be addressed by returning back to the development phase and revising the artifacts, otherwise flaws are to be highlighted so that future works can address them.

2.1.5 Conclusion

The conclusion stage marks the end of a DSR and the results are written up. This works written part is composed of this thesis as well as all written source code and documentation.

2.2 Modeling Languages

This thesis discusses modeling syntaxes and how they can be handled in the COUCHEDIT architecture. To this end the Petrinets and Statecharts syntaxes are introduced here. They will serve as examples to demonstrate problems and applicability of the developed artifact.

2.2.1 Petrinets Syntax

The first modeling syntax introduced are Petrinets. Petrinets are easy to understand and provide simple abstract and concrete syntax. Thus they serve as a suitable tool for explaining concepts in the following chapters. Figure 2.1 shows the abstract syntax metamodel of the used Petrinets syntax. The Petrinets conceptual representation consists of places and transitions. Places also have a token count. each place can have a variable amount of incoming and outgoing transitions, while transitions can have any amount of incoming and outgoing places.

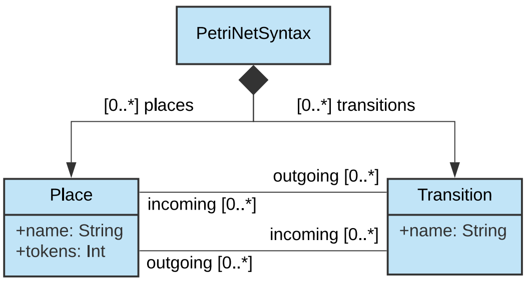


Figure 2.1: Metamodel for a simple Petrinet abstract syntax

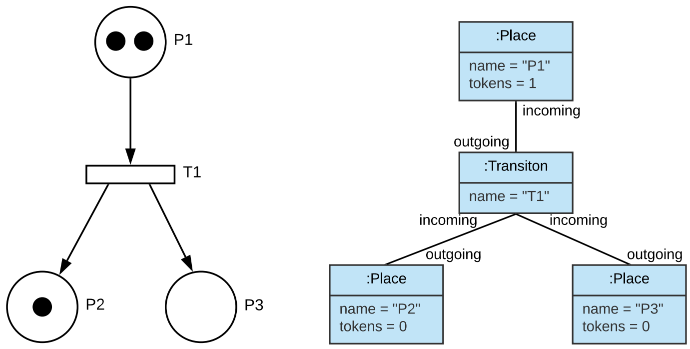
The graphic primitives, a concrete representation of Petrinets is composed of are listed in Table 2.1. places are usually represented as circles, while transitions are depicted as slender rectangles. The number of small black circles inside a place represent this places token count. Furthermore places and transitions possess a label close to their bounding box, that determines their name. Lastly places and transitions have directed connection lines between them. Each connection points from a place or transition towards an element of the opposite type and represents an outgoing connection for the source element and an incoming connection for the target element. A simple concrete instantiation and its corresponding abstract representation is shown in Figure 2.2.

2.2.2 Statechart Syntax

as a second example, Statecharts were chosen. The term Statecharts was first coined in 1987 by D. Harel [4]. Statecharts are an extension of state machines. Their primary feature is

Place	Transition	Token	Label	Connection
			P1	

Table 2.1: graphic primitives used to describe Petrinets



(a) concrete syntax (b) abstract syntax

Figure 2.2: concrete and abstract representation for a simple Petrinet example

that each state can have its own sub state machines. Figure 2.3 describes the abstract syntax metamodel that is used in this thesis. This does not represent a complete specification of the complete Statecharts syntax, rather a sub set which can be used to highlight certain areas of the implementation. Statecharts define a more complex syntax than the aforementioned Petrinets and thus will be used to show the developed artifacts applicability towards more complex modeling languages.

A Statechart is composed of *StateElements* and *Transitions*. StateElements might be connected by Transitions. Each Transition has exactly one source StateElement and one target StateElement. StateElements can either be *PseudoStates* (e.g. initial state, choice, etc.) or normal *States*. A State can either possess a name, a name and a sub state machine, or multiple *Regions*. Each Region contains one state machine. Regions are marked by splitting up States using dashed lines.

2.3 CouchEdit

As noted in Section 1.1, the CouchEdit framework is based on the ideas of Van Tendeloo, et al. [2]. In their paper, the authors criticize that classic approaches to modeling frontends usually create one monolithic construct, which handles both the graphical model display as

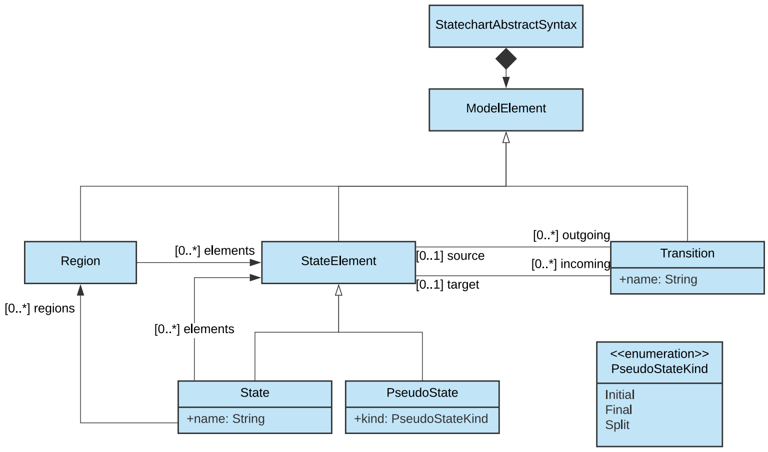


Figure 2.3: Abstract Metamodel of Statecharts

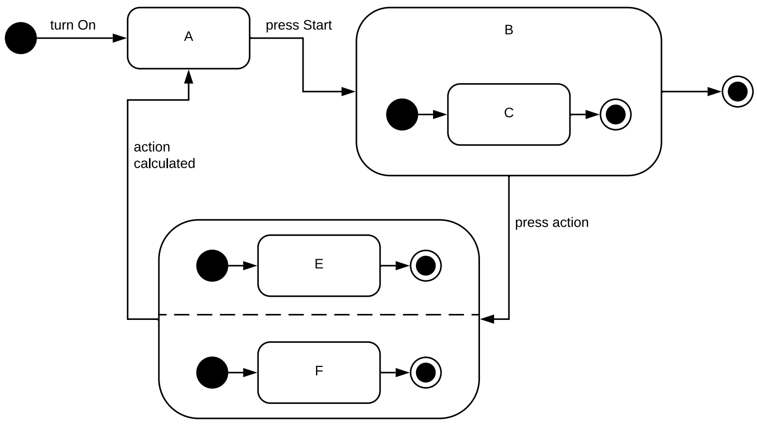


Figure 2.4: Example Statechart

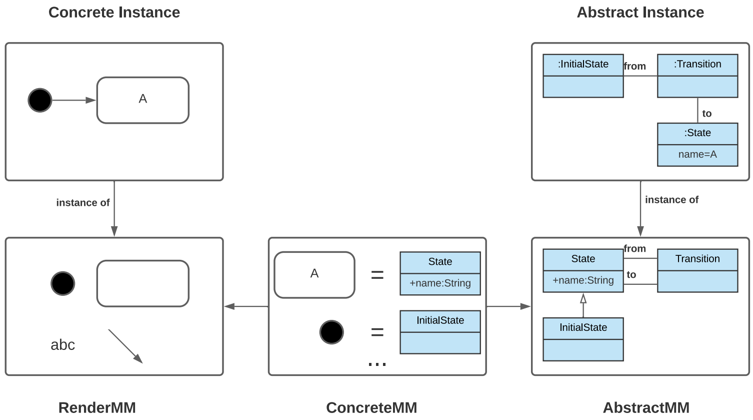


Figure 2.5: Depiction of CouchEdit's approach to separating graphical and abstract representation

well as the corresponding abstract syntax. This creates tight coupling between the two, which causes low flexibility of the UI and high effort requirements, to implement new features.

They propose to solve this, by defining two metamodels, one describing the graphical syntax and one the abstract syntax. The render syntax metamodel specifies graphic primitives that a given graphical modeling syntax is made of. instances of this metamodel are managed by the frontend and a user can manipulate this concrete representation directly by adding, deleting or modifying Elements in the UI (if the frontend allows changes). On The other side stands the abstract syntax metamodel, it describes the conceptual representation of a modeling syntax. Instances of this metamodel are managed by the backend. To then connect these two metamodels, a further metamodel, called concrete syntax metamodel, can be defined. This concrete syntax metamodel specifies a connection that maps the concrete graphical representation onto an abstract representation (Figure 2.5).

Following this architecture, CouchEdit separates frontend and abstract syntax model. The frontend at its core is a simple vector drawing application that implements further functionality for diagrams. A Backend then has the responsibility to analyze these vector graphics and map them onto a corresponding abstract syntax representation. For this, the backend utilizes a set of independent components, that each are responsible for detecting different aspects of the diagram. Each Component contains its own state. When a component makes changes to its state, it also publishes these changes so that other components, interested in these changes, can use them to calculated follow up changes.

This architecture, has the advantage of being very detangled. this allows for components to be improved in isolation from the rest of the system and the clear separation of concerns makes it easier to understand, which component is responsible for what task. Furthermore, components can be swapped out depending on what tasks are needed for the configured modeling syntax.

2.3.1 Data Model

CouchEdit uses a hypergraph of Elements and Relations to represent its data. Each component has its own instance of this graph that only contains the area of the graph, that is interesting to this component. A component can manipulate its personal graph and make these changes known to the system, so that other components can use them as well.

2.3.1.1 Elements

`Elements` are the base type of CouchEdit's hypergraph. Every type of node in this graph inherits from the *Element* type. As all hypergraph instances should be independent from each other, it is not possible to access Elements via their object reference. Thus every Element has a unique id that can be used to reliably identify it across multiple graphs. Elements also have a *probability* attribute that denotes how probable it is that this Element is the result of a correct interpretation of the given hypergraph. This probability object can be set to *Explicit*, which means that, either this Element describes a factual truth or it was marked by the user as not to be changed. Furthermore, Element subtypes can introduce further attributes that contain important information for the given type.

2.3.1.2 Relations

Relations represent edges of the hypergraph. Each Relation connects a set of source Elements to a set of target Elements. They also inherit from the Element type and thus can recursively be connected with relations themselves. The connected vertices are referenced using the *ElementReference* class. This class holds an Elements unique id as well as its type, which then can be used to find a referenced Element in the hypergraph. Figure 2.6 shows the implementation of Relations.

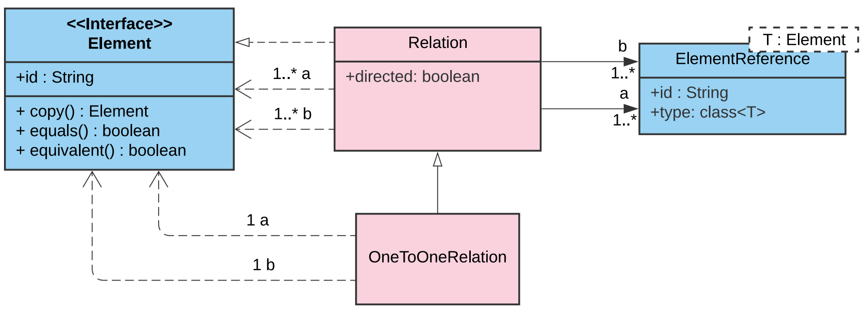


Figure 2.6: Class diagram of Relations

While Relations can connect multiple source and target vertices, the most common Relation type is the *OneToOneRelation*, it describes a connection between exactly one source and one target Element. A common example, for such a OneToOneRelation, is the *Contains* relations, which defines that one Element encloses another.

2.3.1.3 Graphic Objects

Graphic objects (GOs) represent atomic elements, that the user interacts with via the Frontend. An important decision, when designing an Editor is the question on how granular singular graphic objects should be. On a basic level, GOs can consist of either pixels or vectors. a pixel based Approach would require a visual recognition preprocessor, that composes the state of pixels in to comprehensible graphic primitives. While such form of image recognition is outside the scope of the COUCHEDIT project, similar projects such as FLEXISKETCH [5] show the application of this approach. On the other hand most modeling approaches with some form of abstract syntax processing, build specialized graphic elements with the purpose of representing a specific element of the abstract syntax. Therefore mapping the concrete syntax to an abstract representation is made easy. But as a result these graphic elements are very strict in their graphical representation.

COUCHEDIT's goal is it to allow users to construct diagrams using elementary, "multi-purpose" graphic objects as well as specialized elements (if required by the modeling language) [1]. To this end, COUCHEDIT uses a render metamodel (RenderMM). It defines all building blocks that are usable in a concrete representation. COUCHEDIT builds its RenderMM upon the Classification of Costagliola, et al. [6]. This classification defines a graphic element on its lowest level as a primitive graphic shape (e.g circle, line) Via sub-classing, these primitive GOs can then be extended to form specialized, *complex graphic objects*.

In his work, The “Physics” of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering [7], Moody specifies eight variables that can be applied to a single graphic primitive. These attributes are separated into factors that have direct influence on the relations between GOs: Position (horizontal and vertical), shape, size, orientation. And attributes that only serve as visual hints and influence a single GO at a time: color, brightness, texture. COUCHEDIT realizes this separation as well. Factors with spatial influence are contained in the direct info of the shape. While purely visual attributes are externalized into *AttributeBags*. An AttributeBag is a separate element that can be attached to GraphicObjects and contains secondary information about the given Object.

2.3.1.4 HotSpotDefinitions

Some Modeling Notations define relevant zones in a graphical notation, which are not directly modeled by a GraphicObject and thus are not directly visible to the user but exist as a result of the graphical composition. For example in [8], attachment points have been identified as such a relevant zone. An attachment point defines an area around a GraphicObject where a connection line can be "attached". This concept is implemented in a more general form through *HotSpotDefinitions*. A HotSpotDefinition is a relation that, similar to GraphicObjects, has a shape and thus can be a target as well as a source of spatial relations.

One of the more primary applications of HotSpotDefinitions is in the form of *CompartmentHotSpotDefinitions*. Compartments split a GraphicObject into multiple logical sub areas. This is used in multiple advanced notations. For example the Statechart syntax discussed in Section 2.2.2 where a State can have multiple Regions marked by dashed lines.

2.3.2 Application Structure

CouchEdit builds around independent components, called *Processors* that can be enabled or disabled as needed in a given use case. To this end, the center piece of CouchEdit's architecture is the *ModificationBusManager*. Every Processor (including the frontend) can be connected to this bus manager (Figure 1.1). if one Processor publishes a change of the hypergraph, the ModificationBusManager propagates this change to all Processors connected. Each Processor can specify which Element types it is interested and will only receive updates for Elements that match one of those types.

To communicate these changes in the system *ModelDiffs* are used. Each Diff represents a adding, change or deletion of an Element from the Hypergraph. This allows Processors to calculate correct graph changes based on the incoming changes, instead of having to reevaluated the complete graph. This was added into system because it was suspected that always reevaluating the complete state would cause a lot of overhead. Especially Processors that have to calculate spatial relations turned out to profit a lot from diff based calculations [1]. But calculating the correct follow up state, based on a set of Diffs turned out to be highly error prone. Thus it was suspected that it could be beneficial to reevaluate the complete graph when processing lighter tasks, such as Abstract syntax processing [1].

2.3.2.1 Core Processors

CouchEdit implements a set of `Processors` that are integral to most modeling syntaxes. These so called core `Processors` currently are:

SpatialAbstractor This `Processor` has the responsibility to calculate how `GOs` are positioned to each other. Possible position `Relations` include: `RightOf`, `BottomOf`, `Intersect`, etc..

ConnectionEndDetector The `ConnectionEndDetector` finds line `GOs` that could represent a connection to another `GO`. it Then adds `ConnectionEnds` relations from the line to the other `GO`. Furthermore, the `ConnectionEnd` relation has the attribute `isEndConnection` that defines, if this relation origins from the ending point of the line.

Containment The `Containment processor`, checks if one `GO` is contained by another. Meaning, if one `GO` completely encompasses another `GO`, the `Processor` adds a `Contains` relation from the surrounding `GO` to the contained one.

2.3.3 Services

`Processors` may sometimes require to gather additional information, such as. For example it is often needed to find all `Elements` that are connected to a given `Element` by a certain `Relation` type. To this End, CouchEdit provides `Services`. `Services` are separate objects that provide a set of functions. A `Service` can be required by multiple different `Processors` and thus must be effectively stateless.

3 Related Work

this chapter provides information about work in the area of graph translation and concrete to abstract syntax mapping. This primarily encompasses Triple Graph Grammars, as well as work in the area of relaxed conformance modeling.

3.1 Triple Graph Grammars

Triple Graph Grammars (TGGs) were first introduced by Schürr in 1994 [9]. They formalize a specification that can be used to define bidirectional translations between different graph languages [10]. As the name implies, TGGs consist of three graphs, often called source, target and correspondence graph. The source and target graphs are distinct graphs that then are connected by the correspondence graph. The correspondence graph is responsible for realizing correspondence relationships between source and target graphs. Production rules are defined for a given graph triple. By applying these rules to a given instance of the source graph, it can be mapped to a corresponding target graph representation.

This triple graph structure is similar to the architecture employed by COUCHEDIT, explained in Section 2.3. Thus TGGs are a possible mechanism to specify configurations for COUCHEDIT. Furthermore, well defined TGGs can generate *forward and backward graph translations*. As COUCHEDIT strives to provide functionality that requires translation from RenderMM to AbstractMM as well as the other direction, the bidirectionality TGGs provide could be beneficial for future implementations. Nonetheless the approach described in the following thesis does not make use of TGGs. TGG production rules are usually defined as a form of graph grammar rules [10]. In the early development stages of this research it was evaluated that the most taxing aspect of translating concrete to abstract syntax is the recognition of correct concrete representations in the graphical representation. This means that realizing the translation from RenderMM to AbstractMM using graph grammar rules is especially verbose on the left hand side of the production rule. the right hand side seemed to always be rather simple. Furthermore [10] notes that most TGG implementations use inefficient graph grammar parser algorithms. As COUCHEDIT has to provide user feedback during the editing process, performance represents an important aspect in production ready implementations of COUCHEDIT. Nonetheless, TGGs could provide benefits for COUCHEDIT and it may be worthwhile to investigate their applicability in future works.

3.2 Making Metamodels Aware of Concrete Syntax

in their work, Making Metamodels Aware of Concrete Syntax [11], F. Fondement and T. Baar argue that, while abstract syntax definitions are standardized, most language specifications keep the concrete syntax informal. To solve this problem, they propose an approach to defining the concrete syntax and how to link it to the abstract representation.

For this, the authors complement every class of the abstract syntax with a corresponding display scheme. This display scheme is compose of two parts, an iconic and a constraining part. The iconic part defines a set of *DisplayClasses*, these DisplayClasses group Graphical Objects together into visual representation object. On the other hand, the constraining part links these DisplayClasses to an abstract syntax element. This link is realized using *DisplayManagers*. A DisplayManager serves as connection between exactly one model element of the abstract syntax and one display object and has the task of syncing the abstract to the concrete representation. An example of this architecture is depicted in Figure 3.1, for a Petrinet place element. The graphical primitives, a place is composed of, are mapped to a place DisplayClass, to build a places iconic part. This iconic representation is then attached to a place model element, using a place DisplayManager.

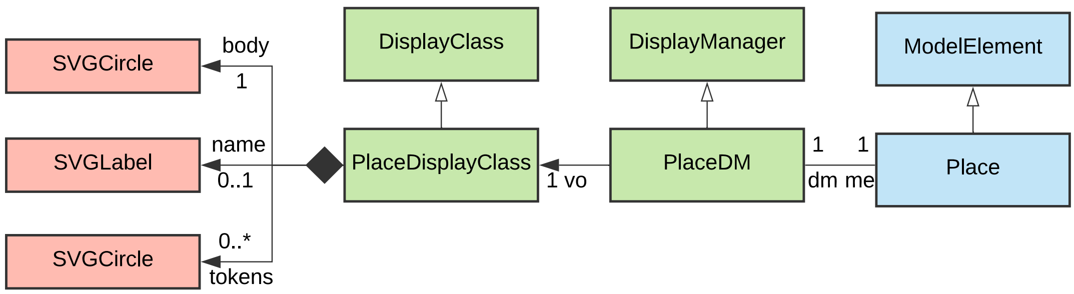


Figure 3.1: Example representation of a Petrinet place element

A DisplayManager has to keep abstract and concrete representation in sync, for this Fondement and Baar utilize OCL invariants, which are defined on the DisplayerManager. For example, an invariant to sync the name of place display schemes could look as following:

```
context PlaceDM
inv: self.me.name->exists() implies
        self.me.name = self.vo.name.text
```

Listing 3.1: OCL Invariant that syncs the name attribute of Place DisplayClasss and model elements.

The Authors keep open, how the mapping from graphical primitives to a display object could be implemented. while the metamodel proposed in this thesis does not introduce an extra layer of abstraction in form of these DisplayClasses, it still is inspired heavily by Fondement's and Baar's work. Especially the constraining part that utilizes DisplayManagers to sync abstract and concrete syntax, served as primary inspiration for the proposed architecture.

In the work "Correctly defined concrete syntax" [12], Baar furthermore provides an algorithm to check syntactic correctness of the concrete representation. While checking the concrete syntax for correctness is out of scope of this work, a future implementation would have to provide the user with syntax checking capabilities, to make sure that the model, designed in the editor is syntactically correct.

3.3 DiaGen

DIAGEN is a project first introduced by Minas and Viehstaedt [13]. It is a system that allows for the specification and generation of modeling editors. DIAGEN supports *free-hand editing*, where the user can manipulate the graphical representation directly. This is the same editing mode that COUCHEDIT supports [1]. In DIAGEN, a diagram consists of a finite set of diagram components [14]. Each component is defined by its attributes. Furthermore, each component has a set of *attachment points*. Attachment points define areas at which different diagram components can be connected. DIAGEN's architecture generates a hypergraph from the visual diagram consisting of these diagram components. This hypergraph is furthermore reduced using a context-free grammar to receive the abstract representation. DIAGEN also supports automatic layouting and user feedback on diagram parts that are not correct under the defined diagram language [14]. DIAGEN provides a textual language as well as a GUI tool, which can be used to specify diagram languages.

While DIAGEN serves as an interesting example for a textual language to define modeling syntaxes, it does not seem applicable to COUCHEDIT. DIAGEN does not create a hypergraph that contains concrete and abstract syntax, instead it parses the concrete syntax to a intermediate hypergraph which is then reduced to the abstract syntax. Furthermore as a code generator, DIAGEN relies on the possibility to implement certain parts of the definition in source code. While the artifact proposed in this thesis does also include a code generator that makes use of this advantage, ultimately it is intended to evolve this work into a runtime interpreter that can not rely on source code implementations.

4 Design Concept

In Section 2.3 it was established that COUCHEDIT is composed of three distinct metamodels as depicted in Figure 2.5. Thus is seems apparent that the proposed concept should also be composed of these three metamodels. The definition of abstract syntax metamodels has already standardized approaches (e.g. Ecore[1]) and thus is mostly ignored. Furthermore, the proposed design currently only supports a render metamodel definition composed of graphic primitives. Complex graphic structures would have to be explored in follow up works. This work focuses on the concrete syntax metamodel, that connects render and abstract syntax and proposes an approach to formalize this connection. The root definition of the COUCHEDIT configuration metamodel is shown in Figure 4.1. throughout the chapter, this root model is gradually being extended with new concepts, until finally all utilized concepts have been introduced (complete metamodel is depicted in Figure A.1).

The primary goal of the proposed design is to hide COUCHEDIT's complexity behind more accessible metamodel definitions. To this end, multiple concepts are employed, that either abstract away from COUCHEDIT's implementation details, or introduce ways to streamline the translation from concrete to abstract representation. Most of these concepts could introduce performance overhead, but as this work focuses on the design aspect, performance analysis and optimization is subsidiary and subject to further research.

4.1 Render Metamodel

COUCHEDIT's render syntax is realized as a set of graphic objects. The framework in its current form primarily focuses on primitive graphic objects, with its class structure being

[1] `https://www.eclipse.org/modeling/emf/`

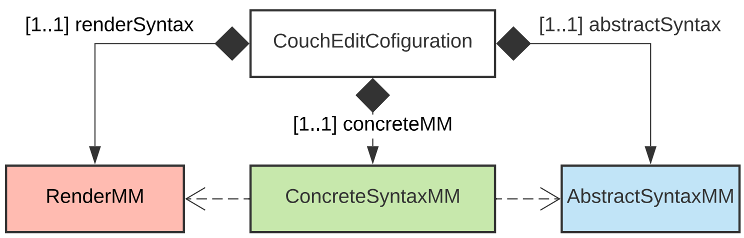


Figure 4.1: Composition of the COUCHEDITConfiguration metamodel

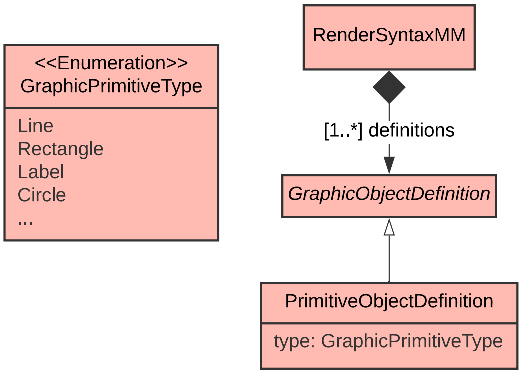


Figure 4.2: Concrete syntax metamodel used in this thesis

derived [1, p.39] from Bottoni and Grau's work, "A Suite of Metamodels as a Basis for a Classification of Visual Languages" [8]. While the notion of complex graphic objects has been introduced in CouchEdit's formal definition, the concept for now only exists in theory. This thesis, because of time constraints, does not explore definition of the concrete syntax all to much and instead focuses only on graphic primitives. The employed render syntax metamodel is describe in Figure 4.2. As shown, it facilitates a basic structure to define graphic primitives needed for the described visual language. Future works would have to explore, how graphic objects could be customized, by integrating attribute bag definitions as well as complex graphic object definitions, to create custom symbols, better suited for certain modeling languages.

4.2 Abstract Syntax

As noted before, the abstract syntax can be defined using existing standards. Nonetheless, the translation metamodel needs to reference the abstract syntax at certain points in this chapter. Therefore, a barebones implementation that is derived from Ecore's own metamodel[2] definition is defined here (Figure 4.3). Notable concepts needed in the definition of an abstract syntax, are foremost definition of classes. This includes, definition of class attributes, as well as concepts such as inheritance and references to other classes. Furthermore definition of simple data types as well as enums is required.

[2] `https://download.eclipse.org/modeling/emf/emf/javadoc/2.9.0/org/eclipse/emf/ecore/package-summary.html`

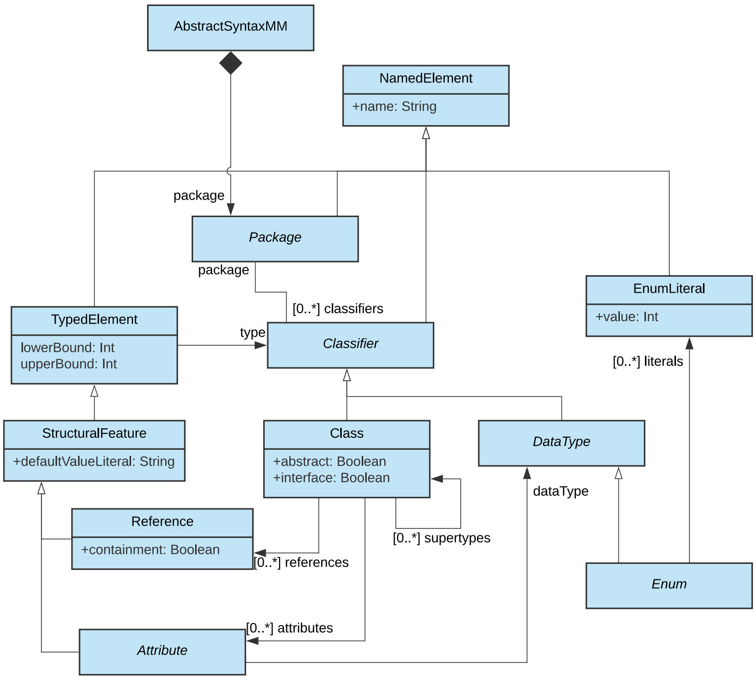


Figure 4.3: Used subset implementation of Ecore

4.3 Language Concepts

the abstract syntax tree usually differs between programing languages. Therefore, when looking at lower level concepts, found in most programming languages (e.g.: arithmetic, branching), their metamodel representation varies while the functionality stays the same. For this reason and because there is already enough documentation on how to specify abstract syntax trees, common language concepts will not be defined, as this depends on the actual language implementation. Instead at points where these language concepts would reside, it is only defined what the given metamodel expects from the implementation and example code snippets are given in an unspecified pseudo language. Despite this, there are certain concepts special to the CouchEdit metamodel, which a language implementation would have to realize. these concepts are presented in the following subsections.

4.3.1 Helper Functions

When defining a translation metamodel, it may become useful to relocate certain procedures or algorithms into separate functions. Reasons for this may be, to access it from multiple points in the definition, to implement a recursive structure or, simply to split up complex parts. Thus the metamodel should provide a facility for the definition of custom functions. That being said, the actual metamodel definition is kept ambiguous, as details depend on the given implementation. In the architecture described here, these helper functions are defined on the highest level of the metamodel and thus are available in all statement blocks, but it would also be feasible to move function definitions into sub scopes, should that provide advantages for certain implementations.

4.3.2 Graph Navigation

Something used a lot in the proposed architecture, is traversal through the hypergraph. It should be possible for the user, to easily get nodes that are adjacent to any given node. In COUCHEDIT's architecture, this is handled by separate services (Section 2.3.3). This means, that the syntax for reaching a node related by a certain relation type looks something like this:

```
val containedGOs = relationService
      .getAllElementsRelatedFrom(go, Contains)
```

Listing 4.1: Example on how to get all `GOs` contained by a given element `go`, in the COUCHEDIT architecture

The `getAllElementsRelatedFrom` function, takes a `GO` and a relation type and returns all `GOs` that have a relation of the given type, that points from the given `GO` to themselves. This and similar helper functions are used repeatedly, thus DSL implementations of the here described architecture should provide some form of abstraction to this service call. In the following code snippets, it is assumed that these functions were masked by member function implementations similar to the following:

```
go.getAllElementsRelatedFrom(Contains)
```

Listing 4.2: The `getAllElementsRelatedFrom` function, implemented as a member function eases navigation through the hypergraph

Implementation in form of such extension functions allows for a more natural navigation trough the hypergraph, which is especially useful when traversing longer distances which will become apparent in the following sections. A list of these navigation functions can be found in Appendix B.

4.4 Syntax Processors

The primary concern when defining a modeling language with clear separation between concrete and abstract syntax, is the question on how to connect these two distinct models. In this point, the designed architecture draws inspiration from Fondement and Baar [11]. The author's proposed idea of connecting abstract und concrete syntax, using `DisplayManagers`,

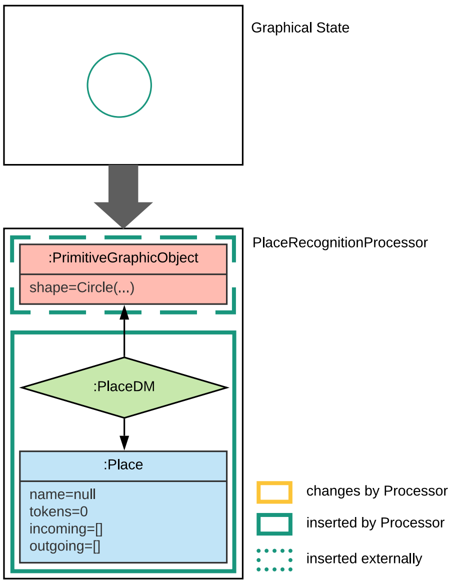


Figure 4.4: PlaceRecognitionProcessor, adding PlaceDM to a GraphicObject that satisfies constraints

serves as a basis that can be built upon. This approach consists of two parts, recognition and synchronization.

4.4.1 Recognition

The recognition part is concerned with detecting patterns in the graphical syntax that have an abstract syntax representation. For this, Fondement and Baar proposed the introduction of a further abstraction layer. This layer composes the graphic primitives and attributes, which represent a model element, into display classes. But the authors keep possible implementation of this abstraction layer open. Furthermore it did not seem suitable to introduce a further abstraction layer, as most processing in the COUCHEDIT architecture is applied directly to the concrete syntax hypergraph, thus mapping graphic objects to a new layer would introduce a set of new problems to be solved. Instead the architecture proposed here, arrived at a different approach. For each type of display manager, a set of constraints can be defined on the hypergraph's graphic objects. Whenever a graphic object satisfies all constraints of a given display manager type it is deemed to be the base element of a concrete representation of this display manager type and an instances of this display manager and corresponding model element are created and connected to the graphic object (Figure 4.4).

These constraints can be defined as simple boolean expressions, that each graphic object in the hypergraph, is checked against. The constraints for a place element, could be defined as follows:

```
go.shape is Circle
go.allRelatedTo(Contains).isEmpty()
```

Listing 4.3: Possible constraints to detect GOs representing a place

These constraints first check if the given GO has the shape of a circle. If that is true, it is also checked if the given GO has any `Contains` relations pointing toward itself. If that is the case, the given circle is contained by another element and thus not clearly identifiable as a place, as it could also possibly represent a token. A corresponding definition for transitions could look as follows:

```
go.shape is Rectangle
```

Listing 4.4: Simple constraint to check for transition representations

These are barebones requirements to identify graphical representations and they could be extended by any amount of further constraints to increase the amount of specificity required from the concrete instance.

4.4.2 Synchronization

With the recognition part in place, it is possible to detect graphical representations and attach corresponding abstract instances. Now the synchronization part is responsible to make sure, abstract and concrete representation reflect the same state. To this end, the metamodel allows for the definition of rules on a `DisplayManager`. This step was explained in detail in [11]. Fondement and Baar propose the usage of OCL invariants as a language for defining these rules. But as the here defined approach does not implement the display class abstraction layer, the concrete representation has to be queried directly. Furthermore, implementing OCL invariants is not necessary and synchronizations could also be handled by assignment statements, which is done in the following. For a place element there are four aspects that have to be synced:

1. Name of the given place
2. Number of Tokens this place has
3. Incoming transitions
4. Outgoing transitions

Reliably determining a place's name poses some special challenges and requires further concepts, introduced later. Determining the token number, on the other hand is easily implemented using the given tools. A simple rule that syncs this attribute could look something like this:

```
dm.me.tokens = dm.go
                    .allRelatedFrom(Contains)
                    .select(go -> go.shape is Circle)
                    .count()
```

Listing 4.5: Rule that syncs the token count of a place element

This statement ensures that the token attribute of the model element is always equal to the number of all GOs with the shape Circle, that are contained by the base graphic object. In a similar fashion, incoming and outgoing transitions can be defined:

```
dm.me.incoming =
      dm.go
          .allRelatedTo(ConnectionEnd,
              rel -> rel.isEndConnection)
          .select(go -> go.shape is Line)
          .collect(go -> go.relatedFrom(ConnectionEnd))
          .select(go -> go.shape is Rectangle)
          .select(go -> go.dm != null)
          .collect(go -> go.dm.me)
```

Listing 4.6: Rule that syncs incoming transitions of a place element

This statement ensures that all transitions, connected to the given DM's GO, by a line, are composed into the list of incoming transitions (Figure 4.5). This exhibits the usual approach to syncing concrete and abstract representation.

4.4.3 metamodel

The resulting sub model for this area of the configuration would look as described in Figure 4.6. The depicted Metamodel allows for the definition of `DisplayManagers`. A `DisplayManager` definition, needs a reference model element type from the abstract syntax metamodel. The corresponding DisplayManager can then be inferred from this element. Furthermore, the `DisplayManager` definition, needs `Constraints` and `Rules`. The `constraints`, are then used by a RecognitionProcessor, to add instances of the inferred `DisplayManager`. On the other hand `Rules` are then used by a DisplayManagerProcessor, to keep model elements aligned. As stated, it would be possible to define `Constraints` and `Rules` as OCL invariants, similar to what Fondement and Baar did. Another possible implementation would be as statement blocks, where `Constraints` return boolean values and `Rules` utilize assignment statements, which would be similar to the here described code snippets. This is not the final iteration of this metamodel as will be shown in the next section.

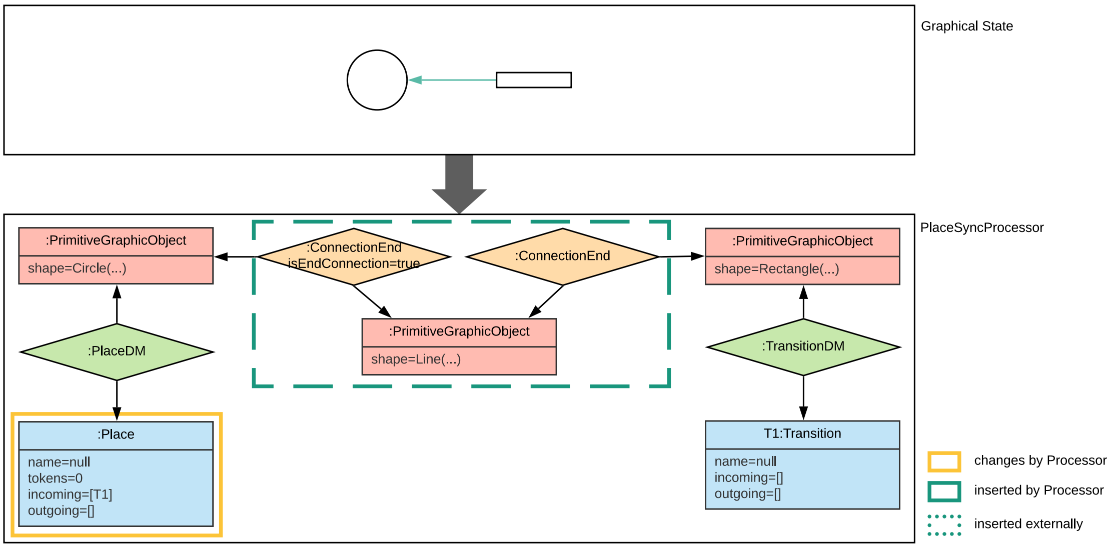


Figure 4.5: A connection line is added to the graph and the PlaceSyncProcessor updates the Place model element

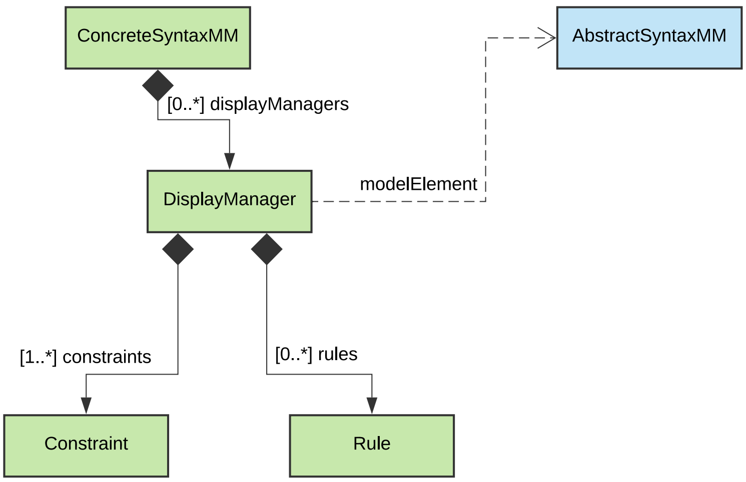


Figure 4.6: First Design of the `DisplayManager` metamodel

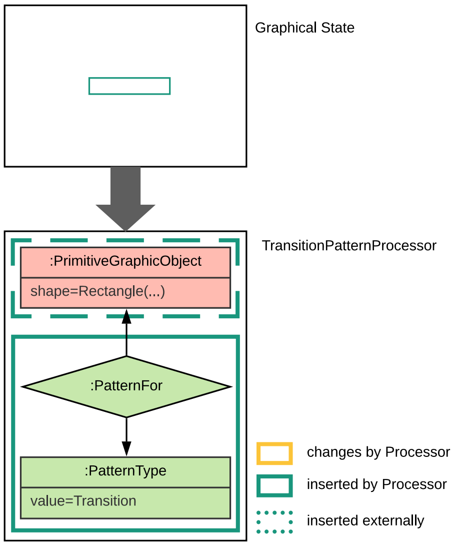


Figure 4.7: `TransitionPatternProcessor` detects a `GO` that satisfies constraints and adds `PatternType`

4.5 Pattern System

When assessing the incoming transition rule (Listing 4.6), it becomes apparent that checking if the connected `GO` represents a transition has to be done manually. In the given example, checking if the `GO` represents a transitions is no big task as transition recognition is handled with only one constraint (Listing 4.4), but when regarding model elements with multiple constraints, this can become a repetitive and inefficient task.

To alleviate this problem, the proposed architecture introduces a Pattern system. A Pattern can mark elements of the graph, as satisfying a certain set of constraints. A `Pattern` is, similar to `AttributeBags`, a separate element in the graph, that is connected to the element it describes, with a relation. It has a single value attribute, that indicates its name. Identical to `DisplayManager` definitions, described in the last section, pattern definitions take a set of constraints. Every Element in the graph is then checked against these `Constraints` and if the Element satisfies all of them, the pattern is added to the given Element. This behavior is depicted in Figure 4.7 for an example transition `Pattern`.

The similarity to the `DisplayManager` recognition part means that it can actually replace this part, which is done in this concept. Instead of checking `Constraints` themselves, `DisplayManager` recognition processors just check if a given GO has a certain pattern type attached and if this is true, the corresponding `DisplayManager` is added. As shown in Figure 4.8, when a `GO`, representing a transition is added, the `TransitionPatternProcessor`, first adds an `PatternType` with the value `Transition`. The `TransitionRecognitionProcessor` then checks the `GO` which now matches the `Transition` pattern and thus adds a `TransitionDM`.

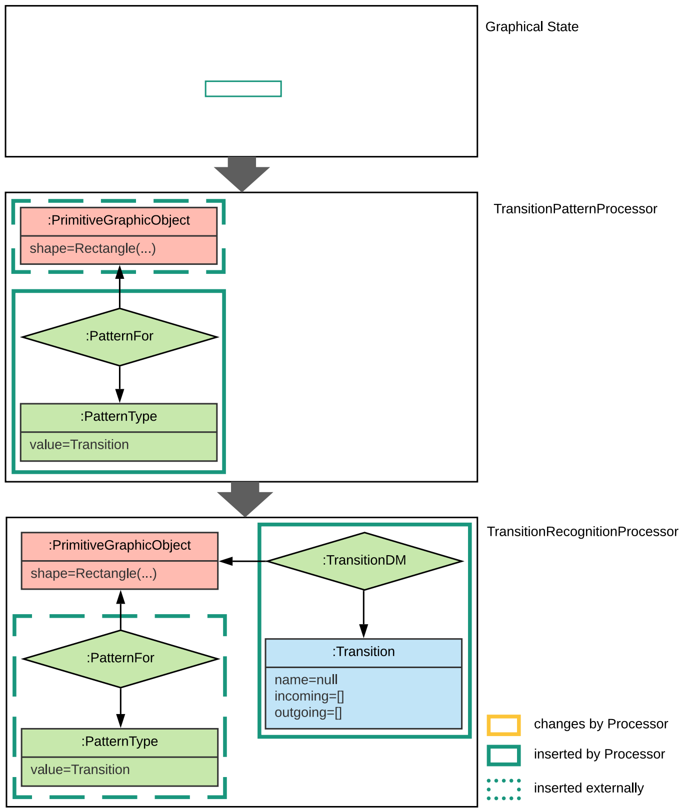


Figure 4.8: `TransitionRecognitionProcessor` adds DM after the `TransitionPattern-Processor` has added a transition `Pattern`

With the definition of a convenience member function that checks if an element has a given pattern, the the rule, syncing incoming transitions, can now be rewritten as follows:

```
dm.me.incoming =
    dm.go
        .allRelatedTo(ConnectionEnd,
            rel -> rel.isEndConnection)
        .select(go -> go.shape is Line)
        .collect(go -> go.relatedFrom(ConnectionEnd))
        .select(go -> go.hasPattern(Transition))
        .collect(go -> go.dm.me)
```

Listing 4.7: Improved incoming transition rule, that also filters for elements with a that match the `Transition` pattern.

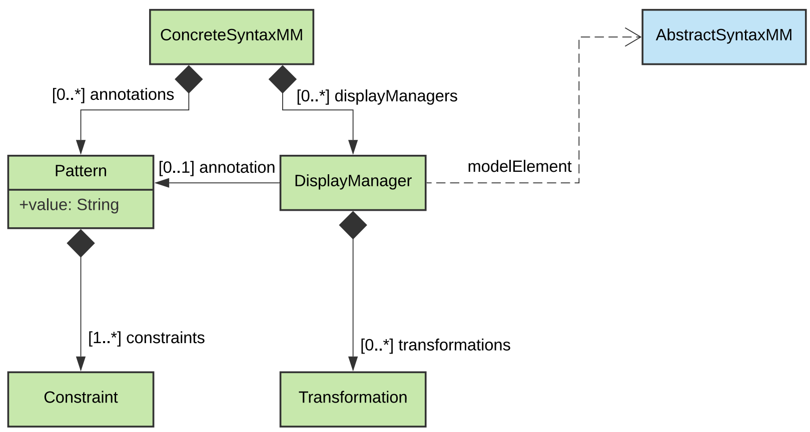


Figure 4.9: Revised SyntaxProcessor metamodel that now includes the Pattern system

It is important to note that the `Pattern` system isn't exclusive to `GOs` that have an abstract representation. For `GOs` with a corresponding `DisplayManager`, the `DM` type can just be checked to find out if it connects a model elements of a certain type. Rather, the pattern system can also be used to recognize all sorts of patterns in the graph, that would otherwise have to be checked repeatedly. For example. it could be used to define a stricter check, on what element in a Petrinet Graph represents a token.

4.5.1 Metamodel

As mentioned in the last section, because of the introduction of the pattern system, the `DisplayManager` definition model has to be revised. Instead of possessing a set of constraints, `DisplayManager` definitions now have a reference to a `Pattern` (Figure 4.9).

4.6 Plugins

While the invariant for incoming transitions, is sufficient for an initial example, it has two major flaws. Firstly, it is not checking if the connecting line has an arrow end. Secondly only lines that are drawn from the transition towards the place, are treated as incoming lines, as noted in Section 2.3.2.1. Fixing these issues in form of a rule would be too verbose for such a common pattern.

For this reason, the plugin system is introduced. Plugins are predefined processors that process specific parts of the hypergraph. These processing areas are not as integral as the ones solved by the core processors and thus are opt-in. Furthermore certain plugin processors can be configured which allows them to cover a wider field of use cases. One example of such a plugin would be the `ConnectionProcessor`. It looks for line `GOs` and checks if they connect two other `GOs`. If that's the case, the processor adds either a `ConnectionTo` relation or a `ConnectionBetween` relation (Figure 4.10), depending on if the line's arrow ends indicate a directed or undirected connection. Adding this plugin allows for a final rewrite of the rule, syncing incoming transitions:

```
dm.me.incoming =
    dm.go
        .allRelatedTo(ConnectionTo)
        .select(go -> go.hasPattern(Transition))
        .collect(go -> go.dm.me)
```

Listing 4.8: Final iteration of the place incoming transition rule

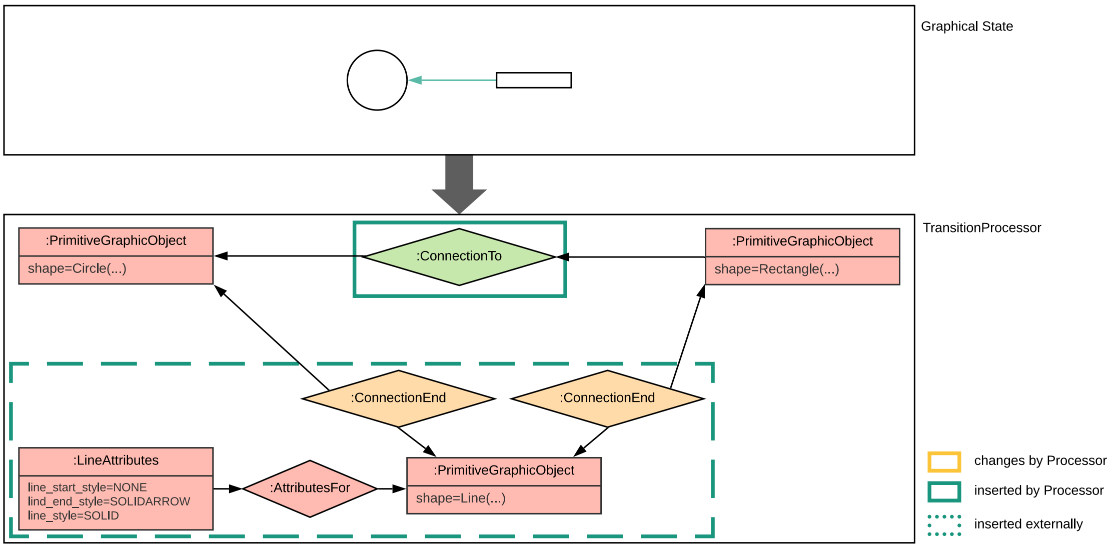


Figure 4.10: `ConnectionProcessor` adds `ConnectionTo` relation on detecting a directed line

4.6.1 Label Processor

Defining a transformation to sync a place's name has been postponed until now, because it poses multiple new challenges. When determining, if a label represents the name of a place object, multiple aspects have to be minded:

1. How close is the label to the place.
2. Are different `GOs` in the proximity that could be associated with this label.
3. Are there other labels that could represent the name of this place.

While possible, checking all these conditions in a syncing statement would require several dozen lines of code. Furthermore it would cause a lot of overhead. For every place, whenever syncing its state, the relation possibility to all labels has to be calculated. If one label is a potential candidate, it also has to be calculated if the label could be a possible candidate for another element. Thus, the idea of building a plugin that handles these calculations arises.

The basic idea of the `LabelProcessor` is, to add `LabelFor` relations from labels to other graphic objects. Furthermore the `probability` attribute is utilized to define how probable it is, that this the intended label association. To be applicable to multiple concrete syntaxes, this plugin poses a new set of requirements:

Requirement 1: Modeling syntaxes require different spatial relations between label and graphic object (label can be contained, surrounding, above/below, etc.).

Requirement 2: Different `GO` shapes require different algorithms to calculate the probability (proximity to a circle cant be calculated the same way as proximity to a rectangle is)

Requirement 3: Not every `GO` needs a label. For example, in the Petrinets syntax, only `GOs` representing either a place or transition, require a label.

Requirement 4: Some label relations can take precedence over others. For example, in Statecharts, a label contained by a simple state, cannot describe a transition trigger event.

4.6.1.1 Label Sub Processor

The `LabelProcessor`, implements a set of sub processors. Each sub processor is responsible for different types of calculations. For example, the `CircleOuterLabel` sub processor calculates the `LabelFor` relation between a circle `GO` and a label `GO` placed somewhere around it. This means, for every combination of primitive shape and spatial relation, a different sub processor is needed. Each sub processor takes an pattern type and a label. It then finds all graphic objects with its designated shape type and filters them by the given pattern type. The resulting `GOs` then receive a `LabelFor` relation from the given label, with a probability calculated by an algorithm implemented in the sub processor.

4.6.1.2 Label ProcessingChain

When regarding requirement 4 it becomes clear that simply enabling the sub processors, needed for a given modeling syntax, isn't enough. A mechanism is needed that allows for different sub processor configurations to influence each other. Therefore the concept of a ProcessorChain was introduced (Figure 4.11). A `ProcessorChain` allows for the configuration of sub processors by introducing an arithmetic like operation system. Processor chains can either be `Operations` or `Terminals`. Terminals represent either a sub processor or nothing. An operation on the other hand defines a link between two Processor chains. Operations possess an `Operator` that defines how the results of the left hand processing chain should affect the results of the right one. For example, the 'and' Operator defines that both processing chains are calculated normally. On the other hand, the 'ifEmpty' operator only calculates the right processor chain, if the left one has returned no values. The 'ifAmbiguous' operator ignores the left chain if more than one relation is returned and instead calculates the right side. Using this ProcessorChain, the label processing for Petrinets could be configured as following:

```
(CircleOuterLabel(Place) and
RectangleOuterLabel(Transition)) ifAmbiguous Nothing
```

This causes the `LabelProcessor` to calculate outer proximity `LabelFor` relations for all circular `GOs` with pattern type `Place` and rectangular`GOs` with pattern type `Transition`. should this calculation return more than one possible relation for a given label, the relations are ignored and nothing is returned instead.

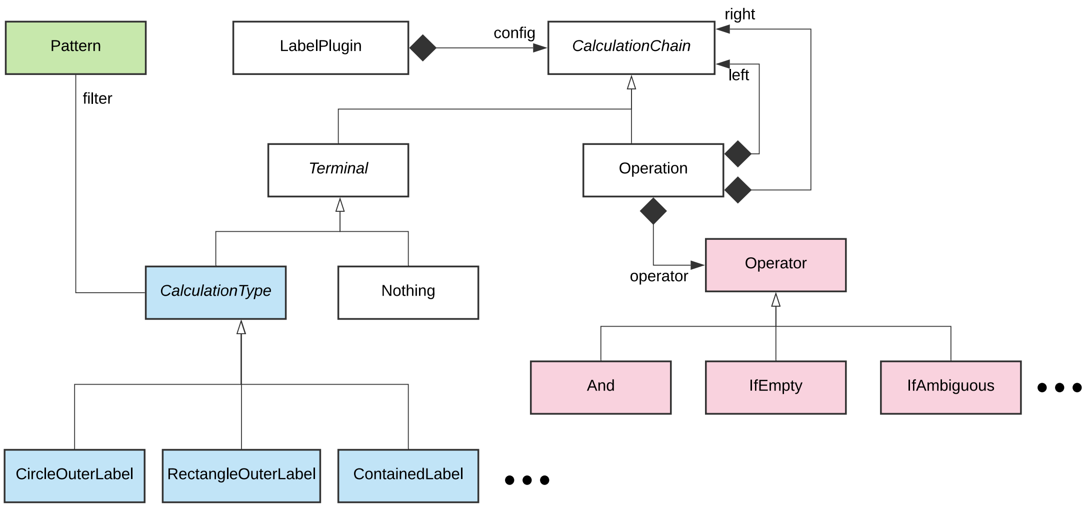


Figure 4.11: Class diagram describing the LabelProcessor's configuration metamodel

This `LabelProcessor` plugin is by no means a perfect implementation for the given problem, rather it provides a general idea of the purpose and applicability for the plugin system, as well as an idea of how specialized the definition of plugin configurations can become. The requirements of plugin configurations, forces DSL implementations of this metamodel to potentially be updated every time a new plugin with custom configuration is added to the system. Alternatively an approach for more general plugin configuration has to be explored

4.6.2 Summary

In this chapter, the developed metamodel was introduced step by step. The base idea for this metamodel was derived from Fondement and Baar [11]. This concept is split into two parts, recognition and synchronization. Recognition is handled by the Pattern mechanism which is responsible for detecting objects in the graphic representation that satisfy a defined set of constraints. The synchronization part is responsible for connecting graphic patterns to a corresponding abstract representation. This system is extended by a plugin system which allows additional predefined processors to be enabled. These plugin processors can require configurations that allow a Plugin to be applicable to a wider range of modeling languages. The complete metamodel is depicted in Figure A.1.

5 Prototype

Along side the development of the metamodel, described in Chapter 4, a prototype was developed as well, which serves as a proof of concepts for design decisions made. This prototype is realized as an extensions of COUCHEDIT's current existing prototype, developed by Nachreiner and presented in [1].

COUCHEDIT's current running prototype is completely written in Kotlin[1], a programing language with the ability to compile to the Java virtual machine platform. Thus, Kotlin enables access to Java's ecosystem of libraries and Frameworks. TornadoFX[2], a JavaFX wrapper, is used as UI Framework. As a modeling tool, GEF[3] was integrated. Dependencies are handled using the dependency injection framework Guice[4], which integrates perfectly with COUCHEDIT's modular architecture. Detailed information about the current prototype and its features are presented in [1].

In theory, integrating COUCHEDIT with a parser that can evaluate language configurations at runtime would be preferable. This would make it possible to hot swap configurations while the software is running. Instead, the prototype was extended by a code generator. This has multiple reasons. The current implementation is relying on the existence of class definitions. Meaning, at certain points in the code object types are checked. A runtime parser would have to either create workarounds for these type checks or compile the needed classes at runtime, using byte code manipulation. A code generator can generate all class definitions directly and thus solves this problem without any extra effort. Furthermore a code generator can work with a metamodel that is not fully defined. As noted in Chapter 4, low level languages concepts were not fully defined, as they depend on the actual language implementation. To circumvent the problem of having to define an actual language implementation for these low level concepts, the implemented parser opts to simply replace these parts with strings. Statements and expressions, can simply be written directly into these Strings, which are then placed into the generated code, without any further analysis of their correctness. This means, that these strings have to contain Kotlin code. Furthermore, needed language concepts (Section 4.3) have to be provided by the Kotlin implementation itself.

The code generator was implemented as a separate sub module of the existing prototype. It contains its own runnable method that, when executed, will generate a new module. This module provides the language configuration, which is loaded by the model application when

[1] `https://kotlinlang.org/`
[2] `https://tornadofx.io/`
[3] `https://www.eclipse.org/gef/`
[4] `https://github.com/google/guice`

started and automatically configures the application for this modeling syntax. For code generation purposes FreeMarker[5] is used. FreeMarker is a language agnostic template engine, and therefore allows for the generation of kotlin code out of the box.

To ease the effort of defining language configurations, it was also decided to implement an internal DSL, using Kotlin's integrated capabilities[6]. Furthermore, Kotlin's script API was utilized, which allows for the relocation of configurations into script files. These files then can be loaded at runtime.

5.1 Kotlin DSL

The COUCHEDIT DSL is composed of three distinct parts, each representing one of the sub metamodels defined by the `CouchEditConfiguration`. Each COUCHEDIT configuration consists of three files, `RenderMM`, `ConcreteMM` and `AbstractMM`, describing each of the three sub metamodels, the COUCHEDIT configuration is composed of.

5.1.1 Render DSL

Given the simplicity of the `RenderMetamodel` defined in this work, the corresponding DSL does not encompass much grammar. The RenderMM DSL provides a set of constants, each representing a shape supported by the framework. A configuration takes a set of these constants. Future work would have to look at how graphical attributes work with the system as well as integration of complex graphic objects. An example of such a `RenderMM` instance for a Petrinets configuration, is shown in Listing C.1.

5.1.2 Abstract DSL

It was noted in Section 4.2 that the `AbstractSyntaxMetamodel` can be defined using existing tools such as Ecore. As developing a functioning Ecore parser was deemed to complex, the prototype instead implements a barebones grammar that allows for the definition of a subset of the Ecore metamodel. Notably the DSL allows for the definition of enums and classes. While interface definitions are worthwhile as well, they are not yet supported.

`StructuralFeatures` allow for the definition of attributes a class possesses. Most Notably they allow for the definition of attributes and references to other classes. A `StructuralFeature` holds a name, as well as upper and lower bound definitions, where an upper bound of '-1' indicates that it is unbound. Furthermore a `StructuralFeature` can hold a default value. The DSL allows for the definition of `StructuralFeatures` in form of three symbols, `attr`, `ref` and `comp`. `attr` is used to define simple attributes, a class definition can hold.

[5] `https://freemarker.apache.org/`

[6] `https://kotlinlang.org/docs/reference/type-safe-builders.html`

Currently supported `AttributeTypes` are: boolean, string, integer, double, unspecified Java classes and objects, as well as any enums defined. `ref` and `comp` can be used to define references to other classes. Because of limitations imposed by the Kotlin DSL, enum as well as class reference names have to be wrapped into string literals. While `ref` defines standard references, `comp` defines compositions. This plays an important role When generating source code from this configuration. References define connections to objects, that can exist by themselves. In the COUCHEDIT architecture that means, they have to exist in the applications hypergraph. Thus they can only be referenced, using `ElementReferences`. On the other hand, composition objects do not exist as separate entities in the graph und therefore must not be referenced using `ElementReferences`. the `AbstractMM` definition for Petrinets is shown in Listing C.2.

5.1.3 Concrete DSL

The `ConcreteMM` definition is the centerpiece of the developed artifact, thus defining this part was main focus during the development process. Similarly to the metamodel defined in Chapter 4, the `ConcreteMM` DSL is composed of three distinct concepts, with an additional possibility of defining helper functions. Each of the concepts is defined in its own scope.

Helpers: The `helpers` scope allows for the definition of `helper` functions. `helper` definitions take a name, multiple arguments, as well as a body. The functions return value is determined by the bodies last statement, as is done in all following statement blocks. Usually Kotlin can infer the type of this value by itself. In the case a recursive function is implemented, the return type has to be specified as Kotlin will run into recursive type checking problems. Because of the nature of a variable argument count, introduced by `param` definitions, body and return type have to be specified using Kotlin's named parameter syntax[7]. To reduce the overhead of implementing functions, parameters as well as the function body have to be written in string literals. Thus these values can then later be inserted into the generated kotlin source code. This enables access the full Kotlin functionality without having to do any extra work. On the flipside, that means the an IDE can not give any syntax checking for code segments encoded into string literals.

Plugins: The `plugins` scope allows available plugins to be enabled. Similar to the `RenderMM` DSL, the plugins scope provides a set of constants, each representing an available plugin. Furthermore, plugins that require additional configuration have their own scope, which provides the grammar required to specify a configuration for this plugin.

Patterns: The `patterns` scope allows for the definition of new pattern types. Each new pattern type is marked by a `def` block which contains the patterns name, followed by a new scope. The scope contains all constraints that a `GO` has to satisfy, to be marked with this pattern. A constraint is by a '+', followed by a string literal. The string literal represents a complete statement block and thus can contain multiple statements, where the last statement represents the return value. These constraint blocks must return a boolean value und thus,

[7] https://kotlinlang.org/docs/reference/functions.html#named-arguments

have to always end in a boolean expression. Furthermore the constraints are completely isolated from each other. This means, casting the `GO` in one constraint will have no effect on following constraints. On the other hand, these constraints are executed in the order they are defined in. If one constraint fails, following constraints will not be checked. this offers the possibility to safely cast a `GO` in all constraints, when a previous constraint has verified the objects type.

DisplayManagers: The `displayManagers` scope allows for the definition of new `DisplayManagers`. Each new manager definition is marked by a `def` block, that contains a string literal specifying the abstract model element and possibly one literal specifying the corresponding pattern. Rules are defined in the same way as pattern constraints are. A rules statement block on the other hand has no specific return type. Instead they directly act on a given `DisplayManager` and manipulate the corresponding model element. Should two different rules manipulate the same attribute, the rule defined last will overwrite changes done by the first one. Rules can be defined in a way as to not change any values of the model element. An Example definition for a Petrinets `ConcreteMM` configuration is shown in Listing C.3.

5.2 Example Configurations

This section exhibits two implementation examples, based on the model syntaxes presented in Section 2.2.

5.2.1 Petrinets

As a first example a Petrinets configuration was implemented. This configuration follows the model defined in Section 2.2.1, while the complete implementation is shown in Appendix C.1. The complete configuration script, composed of RenderMM, AbstractMM and ConcreteMM, is 106 lines long[8]. The Kotlin code generated by this configuration encompasses 708 lines.

Shapes configured in the RenderMM (Listing C.1) are, circle, rectangle, line and label. With future development of a more sophisticated RenderMM configuration, the circle shape could be defined separately, once as a small black circle and once as a bigger circle without fill color.

The AbstractMM (Listing C.2) definition has two Class definitions. The Place Class has one attribute name of type string, another attribute tokens, with type integer, a fixed bound to "1" and a default value "0". Furthermore the Place Class has two references, incoming and outgoing, with the type Transition and an unlimited upper bound, indicated by "-1". The Transition Class is defined in the same fashion.

The ConcreteMM (Listing C.3) definition has the connection and label plugin, mentioned in Section 4.6 enabled. The connection plugin in its current iteration does not take a

[8]These metrics were counted with the cloc tool: `https://github.com/AlDanial/cloc`

configuration. The label plugin on the other hand is configured to detect labels that surround a graphic object with the pattern Place or Transition. If this results in more than one possible LabelFor relation, the result is discarded as noted by "`ifAmbiguous nothing`". One `helper` was defined, which finds the text, a place or transition is labeled with. The `patterns` scope has three patterns defined, `PLace`, `Token` and `Transition`. `Place` and `Transition` constraints are defined similar to the examples given in Section 4.4.1, but the `Transition` pattern contains an additional constraint that checks if a rectangle has one side at least double the length of the other side. The `Token` pattern is used to detect elements that actually represent a token, to ease the process of syncing the token count. The `DisplayManagers` scope has two definitions, one `PlaceDM` and one `TransitionDM`. The `PlaceDM` connects `Place` pattern and `Place` model element.

The Petrinets configuration provides an example on how to use the developed prototype to define COUCHEDIT configurations.

5.2.2 Statecharts

as a second example, the Statecharts syntax defined in Section 2.2.2 was realized. But the configuration does leave out some of the described functionality. This was done to align this implementation closer with the example presented in [1]. therefore this configuration can then later serve as a point of evaluation. Notably, the implementation of PseudoStates was left open. The complete source code of this implementation is listed in Appendix C.2. While the Petrinets example was explained in detail to demonstrate the usage of the developed DSL, this section only describes noteworthy aspects of the implementation.

Similar to the Petrinets implementation presented in the prior section, this configuration has the LabelPlugin and ConnectionPlugin enabled. Furthermore the *CompartmentDetectorPlugin* is enabled which enables compartmentalization as described in Section 2.3.1.4.

This implementation showcases how patterns can be used to split up problems into smaller parts. A State can have any amount of Regions to solve this the CompartmentDetectorPlugin is employed which recognizes if an area is split up into multiple parts. It is not only possible to split a State into two regions, it is also possible to further split up one Region into two. This means that a CompartmentHotSpotDefinition itself can have two more CompartmentHotSpotDefinitions. Therefore it is important to find the deepest CompartmentHSDs as only they represent a Region. To solve this problem the implementation makes use of two different patterns, `DashedRegion` and `Region`. The `DashedRegion` pattern is responsible for detecting if a CompartmentHSD is created through a *dashed* line and is a compartment of a State object. The `Region` pattern is then responsible for checking if a CompartmentHSD that satisfies the `DashedRegion` pattern has no further sub compartments, making it a Region of the parent State.

In the initial design labeling plugin was configured to prioritize States over Transitions for labeling. This was deliberately designed to solve problems such as shown in Figure 5.1. While this concrete representation can easily be interpreted by a human, without the a

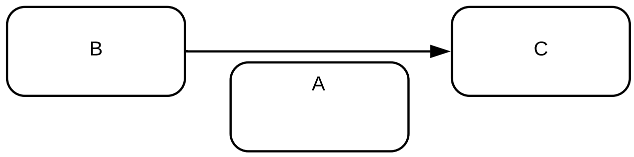


Figure 5.1: Example of problematic labeling without precedence rules.

precedence rule, it would become problematic for the LabelPlugin to correctly assign the label. But this precedence rule turns out to be problematic when handling sub state machines. With this rule it becomes impossible to label Transitions that are part of a sub state machine. The State precedence rule will block all transition labeling for sub states. This makes the rule not applicable to solve the given problem and thus mostly invalidates the usefulness of the LabelPlugin for the given problem. Instead the LabelPlugin has to be configured so that it checks State and Transition labels independently from each other. This means to solve the problem depicted in Figure 5.1 LabelFor relations have to be checked manually. To this end the helper `hasSameParent` is defined. This function check if two elements are contained by the same parent or no parent at all. if this is true, both elements are part of the same hierarchy. This function is used to check if label and Transition are part of the same state machine and thus could be associated to each other. This provides a functional workaround to solve the given problem. But the drawback of this is that most of the functionality the LabelPlugin intends to provide cannot be used. Instead, when finding a corresponding label, all LabelFor relations have to be scanned and filtered in the SyncProcessors increases complexity of the implementation.

6 Evaluation

One of the most important parts of a Design Science Research process is the evaluation. The developed artifact has to be analyzed and it has to be established, how well the defined goals are realized by the artifact.

The main criteria that has to be evaluated, is the applicability of the developed artifact. Meaning, it has to be analyzed how well the developed artifact satisfies the defined goal. To this end it seems worthwhile to evaluate the prototypes performance as well as its developer usability.

6.1 Performance

To provide performance optimization possibilities, CouchEdit was developed with the *Diff* system in mind. Diffs provide the possibility to calculate changes, without the need for reevaluating the complete hypergraph. On the flipside, this means that all possible states of the graph have to be minded. Using this Diff based approach was evaluated as laborious and error prone, but proved invaluable to improve performance of certain processing tasks [1]. Nachreiner's test results showed that especially language specific processing tasks only took up a small amount of the complete processing time. On basis of this result it was decided, that the developed artifact, primarily concerned with language processing, can produce components that reevaluate the complete graph on every change. Plugin processors, such as the LabelProcessor that is expected to cause high load, are still implemented on application level and thus can make use of the performance benefits granted by the Diff system.

While this research can not produce thorough performance results, it was nonetheless attempted to provide a simple indication of the artifact's impact on performance. To this end the `StateGridTest` defined in [1] (Appendix D) was adapted to work with the code generated by the Statecharts implementation (Section 5.2.2). The StateGridTest was then run 50 times for both the reference implementation of [1] and the implementation of Section 5.2.2. The tests were carried out on a modern system containing an Intel Xeon E2-1231 v3 CPU with 3.40GHz clock an 16GB DDR3 RAM with 1600MHz clock. the system was running windows 10 version 2004. The results of this test suit are depicted in Figure 6.1. It becomes immediately clear that the generated code performs worse than the reference implementation. especially with higher counts of GraphicObjects the generated implementation needs exponentially more time then the reference values. During test runs it became clear that the generated implementation starts to max out CPU capacity very early on. Thus the rapid growth in processing time can be attributed to hardware bottlenecks.

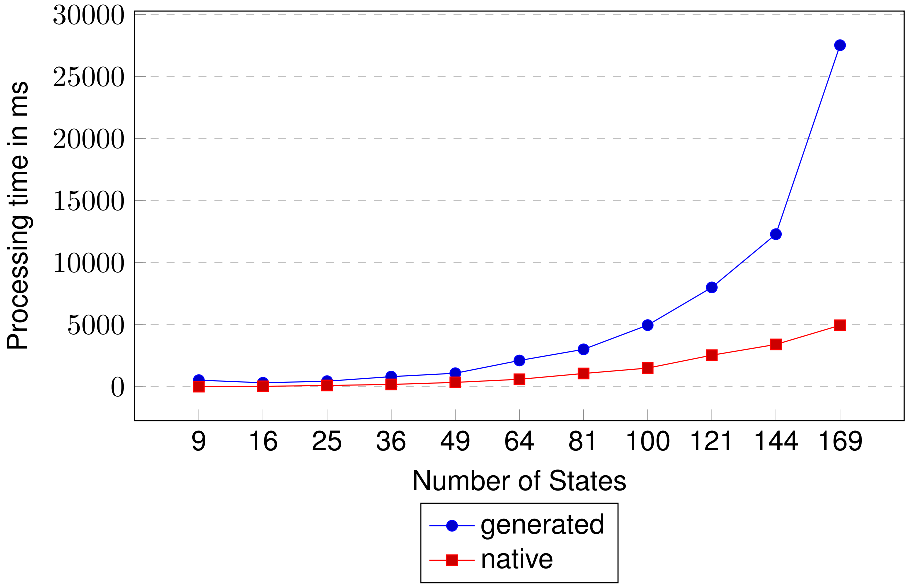


Figure 6.1: Performance comparison between reference implementation and generated implementation in the StateGridTest

While the exact culprit responsible for this rapid growth of processing time is not immediately clear, there are still multiple points that could attribute to this bad performance. First an foremost reevaluating the complete hypergraph in the current implementation is probably the biggest contributor to processing time. In the undirected architecture of CouchEdit, publishing a single change to the hypergraph means that every processor interested in this change will be triggered. The current implementation of the code generator is not able to prune the elements a processor is interested in. This would require analysis of constraints and rules, which requires a complete metamodel definition of this part. Therefore every change published triggers all processors no matter if they are interested in the type of change. This, combined with the fact that each generated processor reevaluates the complete state causes a lot of overhead. Future implementations could try to amend this problem by introducing the aforementioned constraint and rule analysis and use it to prune the hypergraph of each processor. Furthermore it could be possible to reintroduce the Diff based architecture into generated processors. This requires that it is possible to determine which parts of the hypergraph are actually affected by a given Diff. lastly, the developed metamodel introduces a more structured ordering of processors. This could be used to split up processors into sub groups. This way processors in the same group will calculate changes until no processor can produce new changes. All changes calculated are then bundled and passed to the next group. This could reduce the number of times a generated processor is triggered and has to reevaluate the state.

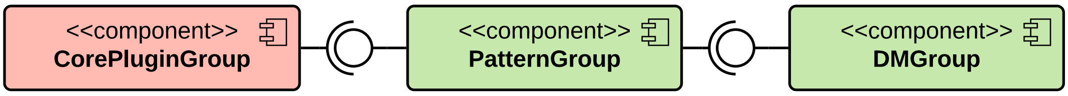


Figure 6.2: `Processors` split into subgroups, where one group is only activated if the previous group has finished processing.

6.2 Usability

Developer usability ask the question on how easy it is to use a tool. The main focus of this research was to develop an artifact that simplifies the process of implementing new modeling syntax configurations for CouchEdit. To this end, it has to be evaluated how well the developed artifact achieves this goal. To find a comprehensive answer to this question, a user study is required. Given the state of the developed artifact as well as the time required to conduct a survey, this goes far beyond the projects scope. Therefore, alternative metrics have to be dissected, in order to determine an inclination regarding this question.

One metric to highlight is conciseness. As an abstraction of the CouchEdit architecture, the developed artifact should be able to implement similar concepts in less lines of code. The example implementations reflect this. The Petrinets implementation is 106 lines long and produces 708 lines of code. Equally, the Statecharts implementation consists of 172 and generates 1256 lines. In both examples, this means, on average of over 6 lines of source code are generated per line of the configuration. Of course, the code generator most likely generates code that is more verbose than an equivalent written by a developer. therefore this metric is flawed and only serves as a suggestion for the possible usability.

To further reinforce claims of conciseness, the Statecharts example (Section 5.2.2) was modeled as closely as possible after the Statecharts configuration implemented in [1]. An exact replica of the modeling syntax implemented by Nachreiner is not possible. The abstract syntax metamodel defined by them, utilizes `Relations`, which the AbstractMM defined in Section 4.2 does not support. This results from the close orientation towards Ecore. Furthermore, the `ConcreteMM` is opinionated in its approach to define configurations and thus, can naturally not provide the same flexibility as an application level implementation has. Nonetheless, the application level implementation is 1900 lines long [1], and was realized here in 172 lines. This is primarily contributed to the fact that implementations using the metamodel do not have to mind every possible state of the graph, as well as the reduction of a lot of boilerplate code.

While the artifact abstracts away from the implementation details of `Processors`, the developer still requires certain knowledge about the framework. Most prominently a developer has to understand the hypergraph. Constraints and rules usually require querying of the hypergraph, therefore it is inevitable to know the existing element types as well as possible Relations and their meaning. Furthermore, it has to be understood which changes, different plugins apply to the graph, as well as what functions are available. Nonetheless, the amount of knowledge required, is still reduced.

6.2.1 plugins

The plugin system proved as a worthwhile addition to the architecture. The example language configurations, described in Section 5.2 show of how complex tasks can be trivialized if the correct plugin is provided. With the plugin system, new `Processors` can be added to the library of existing processors, should problems arise that are difficult to solve within the generalized metamodel architecture. Plugins are developed on the application level an thus have access to COUCHEDIT's complete architecture. On the other hand, this means that developing plugins requires knowledge over the complete COUCHEDIT framework, which the developed artifact is trying to abstract. Therefore plugins should be implemented as general as possible, so that they can be reused as much as possible. Therein lies the challenge of the plugin system. The usefulness of a plugin depends on how well it can be adapted to different modeling syntaxes. This means that plugins have to be developed with great care. In the example Statecharts implementation (Section 5.2.2), the `LabelProcessor` prototype failed to solve all labeling requirements imposed by the syntax. This demonstrates the volatility of hastily developed plugin. On the other hand, the ConnectionPlugin, thanks to its simplicity was able to provide a concise solution for the given problem.

6.3 Discussion of DisplayClasses

During initial evaluation of the architecture proposed by Fondement and Baar [11] it was decided to not adapt the concept of `DisplayClasses`, proposed by the Authors. This decision was made, because their implementation details were not certain. It was estimated, that `DisplayClasses` would introduce further complexity without providing any significant advantages. with regards to the developed artifact, this turned out as mostly true. Nonetheless, there are certain scenarios as well as considerations towards future work that could make an implementation utilizing `DisplayClasses` relevant.

When tightening requirements towards the concrete syntax DisplayClasses could prove useful. For example, if it is required that a Place in the Petrinets syntax always has a name, the same part of the graph has to be queried two times. First the PlacePatternProcessor would have to check if a label is assigned to a Place before it can add the Pattern element of type Place. Then the PlaceRecognitionProcessor can add the PlaceDM. Afterwards the PlaceSyncProcessor has to query the same part of the Graph that the PlacePatternProcessor already queried, to find the corresponding label. Using DisplayClasses, this double querying could be prevented, as the PlaceSyncProcessor would have to just get the Label that is connected to the DisplayClass already, similar to the invariant shown in Listing 3.1.

Another scenario that could make `DisplayClasses` relevant is concerned with syntax feedback in the frontend. DIAGEN marks graphical element in a different color, if they are part of a syntactic correct concrete representation [14]. Such a feature could be conceivable in future iterations of frontends for COUCHEDIT. With the pattern system developed in this thesis, there does not exist a direct connection between all `GOs` and the abstract representation they are part of. Using `DisplayClasses`, every `GO` has a direct connection to the abstract representation they are mapped to. Thus, `DisplayClasses` would trivialize implementation of a feature similar to DIAGEN's.

7 Conclusion

This work presented an experimental architecture that can be used to define modeling syntax configurations for COUCHEDIT. To this end, the approach of Fondement and Baar [11] was adapted and extended to fit the needs of the COUCHEDIT architecture. Their approach is composed of two parts, recognition and synchronization. These two mechanisms where integrated into the developed artifact as the core concept. This concept was then extended by a plugin system, which allows for the addition of modular components which can be configured to satisfy requirements of different modeling syntaxes.

It was shown, that the developed artifact seems applicable for the task of specifying modeling syntaxes for COUCHEDIT. But while it is possible to define modeling syntaxes, especially the definition of well designed Plugins is key to general applicability and developer experience. Furthermore, performance test results indicate that the developed artifact in its current iteration produces configurations that increase processing overhead tenfold compared to a diff based implementation on application level.

7.1 Future Work

This research poses a set of new questions that have to be answered. First and foremost it has to be evaluated if the systems performance can be improved to reach acceptable levels. Towards this goal there are multiple mechanisms that could be researched.

One of the next major iterations of this artifact would be to develop a complete DSL implementation. This would for multiple optimizations. First, the hypergraph each processor requires could be pruned to only contain the elements that are relevant to a given processor. in its current iteration the code generator has no information about the defined constraints and rules, this means that each processor has to be configured to potentially use all possible elements of the hypergraph. therefore processors are triggered even when changes are published that have no importance for them. Furthermore it could be researched if the extra information given by a complete metamodel definition can be utilized to generate processors using ModelDiffs. The results provided in Section 6.1 indicate that evaluation of the complete hypergraph does not scale well. If it is possible to reevaluate only parts of the hypergraph that are actually affected by incoming changes, this could improve performance greatly.

The plugin system proved to be a critical mechanism for the applicability of the developed metamodel. They can take up many tasks from the user and bring further performance improvements as they are implemented on application level an can make use of the available optimization schemes. But Section 5.2.2 showed that the design of a plugin is crucial to ensure applicability a wide range of modeling syntaxes. Thus it has to be evaluated which plugins are needed and how they could be implemented. The LabelPlugin was shown of as a Plugin that is integral to many Modeling languages and thus has to receive further attention to reach its full potential. Moreover, the CompartmentPlugin as well as the ConnectionPlugin could receive further improvements by introducing configuration possibilities for them. Besides the existing Plugins it also seems worthwhile to further investigate mechanisms that could be implemented in the form Plugins. To this end the Works of [2] and [6] could provide a basis to build upon.

It is also important to investigate the artifacts applicability towards correctness checking. It is critical for editor usability to convey to the user if the designed concrete model can be mapped to a correct abstract representation. To this end if was evaluated that the introduction of DisplayClasses could result in a clearer connection between abstract an concrete syntax and thus make it easy to decide which graphic objects art part of an abstract representation. Additionally the work of Baar [12] could be used as a foundation to introduce correctness checking into the framework.

Distributed modeling is one of the potential application of CouchEdit. Meaning development using multiple frontends with only a shared abstract representation. For this to be possible the architecture has to be able to translate from abstract to concrete syntax. in its current iteration the architecture is not able to do this. The artifact generates processors that strictly translate from concrete to abstract representation. Therefore it has to be evaluated if it is possible to introduce this bidirectionality into the prototype. if turns out to be problematic reevaluations using a TGG approach could help to amend these Problems.

An important part theorized for the COUCHEDIT framework is the general model action mechanism. In the form of Suggestions they can provide the user with options that can be used to disambiguate the hypergraph or provide quick fixes that give the user fast actions to transform a malformed graph into something correct. The developed artifact currently does not support any functionality for model actions. To this end further research is needed to find out how this functionality could integrate with the system.

In the early stages of this research it was decided to use the approach proposed by Fondement and Baar [11]. Developing an alternative architecture that builds upon TGGs could bring multiple benefits. Especially the bidirectionality would bring advantages in regards to the future of this project While the incremental

A CouchEdit Configuration Metamodel

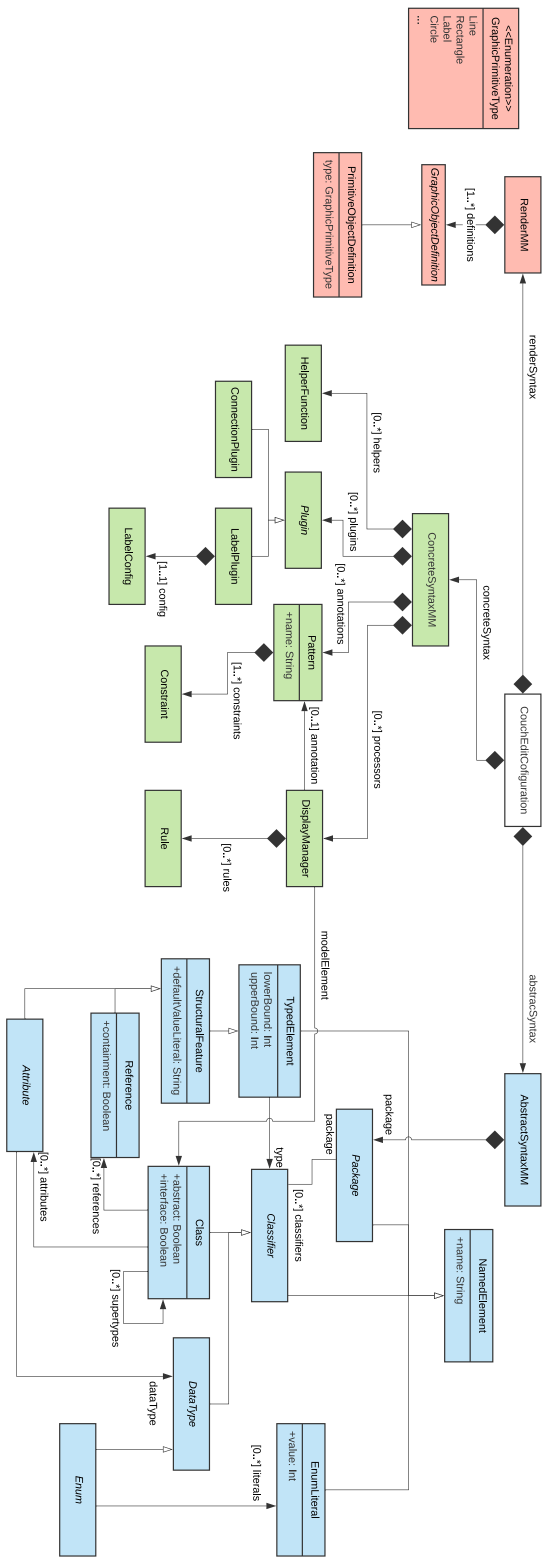


Figure A.1: Overview of the complete designed metamodel

B Navigation functions

To ease traversal of the hypergraph, a DSL implementation of the COUCHEDIT configuration should provide abstractions for the following Service methods:

- tbd

C Kotlin DSL Configurations

C.1 Petrinets Configuration

Render Metamodel

```
package petrinets

import de.uulm.se.couchedit.dsl.dsl.concretesyntax.renderMM

renderMM {
  circle
  rectangle
  line
  label
}
```

Listing C.1: `RenderMM` definition for Petrinets

Abstract Metamodel

```
package petrinets

import de.uulm.se.couchedit.dsl.dsl.abstractsyntax.abstractMM

abstractMM {
  Class("Place") {
    attr("name", STRING)
    attr("tokens", INT, 1, 1, 0)
    ref("incoming", "Transition", -1)
    ref("outgoing", "Transition", -1)
  }

  Class("Transition") {
```

```
    attr("name", STRING)
    ref("incoming", "Place", -1)
    ref("outgoing", "Place", -1)
  }
}
```

Listing C.2: `AbstractMM` definition for Petrinets

Concrete Metamodel

```
package petrinets

import de.uulm.se.couchedit.dsl.dsl.translation.concreteMM

concreteMM {

  plugins {
    connection
    label {
      (circleOuterLabel("Place") and
          rectangleOuterLabel("Transition")) ifAmbiguous
          nothing
    }
  }

  patterns {
    def("Place") {
      +"""go.shapeClass == Circle::class.java"""
      +"""go.allRelatedTo(Contains::class.java).isEmpty()"""
    }

    def("Token") {
      +"""go.allRelatedTo(Contains::class.java).filterPattern
         ("Place").size == 1"""
      +"""go.allRelatedFrom(Contains::class.java).isEmpty()
         """
      +"""go.shapeClass == Circle::class.java"""
    }

    def("Transition") {
      +"""go.shapeClass == Rectangle::class.java"""
```

```
      +"""(self as GraphicObject<Rectangle>).shape.run { h *
         2 < w || w * 2 < h }"""
    }
  }

  displayManagers {
    def("Place", "Place") {
      +"""
      dm.me.tokens = dm.go
          .allRelatedFrom(Contains::class.java)
          .filterPattern("Token")
          .count()
      """

      +"""
      dm.me.name = dm.go
          .allToRelations(LabelFor::class.java)
          .filter { it.isUnambiguous() }
          .takeIf { it.size == 1 }
          ?.first()
          ?.getA()
          ?.shape
          ?.text
      """

      +"""
      dm.me.incoming = dm.go
           .allRelatedTo(ConnectionTo::class.java)
           .filterPattern("Transition")
           .mapNotNull { it.dm<Transition>()?.me?.ref() }
      """

      +"""
      dm.me.outgoing = dm.go!!
          .allRelatedFrom(ConnectionTo::class.java)
          .filterPattern("Transition")
          .mapNotNull { it.dm<Transition>()?.me?.ref() }
      """
    }

    def("Transition", "Transition") {
      +"""
      dm.me.name = dm.go!!
```

```
            .allToRelations<LabelFor>()
            .filter { it.isUnambiguous() }
            .takeIf { it.size == 1 }
            ?.first()
            ?.getA()
            ?.shape
            ?.text
        """

        +"""
        dm.me.incoming = dm.go
            .allRelatedTo(ConnectionTo::class.java)
            .filterPattern("Place")
            .mapNotNull { it.dm<Place>()?.me?.ref() }
        """

        +"""
        dm.me.outgoing = dm.go
            .allRelatedFrom(ConnectionTo::class.java)
            .filterPattern("Place")
            .mapNotNull { it.dm<Place>()?.me?.ref() }
        """
    }
  }
}
```

Listing C.3: `ConcreteMM` definition for Petrinets

C.2 Statecharts Configuration

Render Metamodel

```
package statechartsnew

import de.uulm.se.couchedit.dsl.dsl.concretesyntax.renderMM

renderMM {
  line
  label
  roundedRectangle
}
```

Listing C.4: `RenderMM` definition for Statecharts

Abstract Metamodel

```
package statechartsnew

import de.uulm.se.couchedit.dsl.dsl.abstractsyntax.abstractMM

abstractMM {

  AbstractClass("ModelElement")

  AbstractClass("StateElement") {
    superTypes("ModelElement")
    ref("incoming", "Transition", -1)
    ref("outgoing", "Transition", -1)
  }

  Class("State") {
    superTypes("StateElement")
    attr("name", STRING)
    ref("regions", "Region", -1)
    ref("elements", "StateElement", -1)
  }

  Class("Region") {
    superTypes("ModelElement")
    ref("elements", "StateElement", -1)
  }

  Class("Transition") {
    superTypes("ModelElement")
    attr("event", STRING)
    ref("source", "StateElement")
    ref("target", "StateElement")
  }
}
```

Listing C.5: `AbstractMM` definition for Statecharts

Concrete Metamodel

```
package statechartsnew

import de.uulm.se.couchedit.dsl.dsl.translation.concreteMM

concreteMM {

  plugins {
    connection
    compartmentDetector
    label {
      (containedLabel("SimpleState")) and
          lineCenterLabel("Transition")
    }
  }

  helpers {
    helper("CompartmentHotSpotDefinition.isDashed", body =
       """
    getBSet()
          .first()
          ?.findAttribute(GraphicAttributeKeys.LINE_STYLE)
          ?.value
          ?.equals(LineStyle.Option.DASHED)
          ?: false
    """)

    helper("haveSameParent", "first: Element?", "second:
       Element?", body = """
      first?.relatedTo(Contains::class.java)?.id ==
      second?.relatedTo(Contains::class.java)?.id
    """)
  }

  patterns {
    def("StateElement") {
      +"""go.hasPattern("State") || go.hasPattern("
         PseudoState")"""
    }

    def("State") {
```

```
      +"""go.shapeClass == RoundedRectangle::class.java"""
    }

    def("SimpleState") {
      +"""go.hasPattern("State")"""
      +"""go.dm<State>()?.me?.elements?.isEmpty() ?: false"""
      +"""go.dm<State>()?.me?.regions?.isEmpty() ?: false"""
    }

    def("NameRegion") {
      +"""go is CompartmentHotSpotDefinition"""
      +"""
        val a = (go as CompartmentHotSpotDefinition).getASet
            ()
        a.size == 1 && a.first().hasPattern("State")
      """
      +"""(go as CompartmentHotSpotDefinition).index.indexUL
          == (0 to 0)"""
    }

    def("BaseRegion") {
      +"""go is CompartmentHotSpotDefinition"""
      +"""(go as CompartmentHotSpotDefinition).getASet().
          filterPattern("State").isNotEmpty()"""
      +"""!(go as CompartmentHotSpotDefinition).isDashed()"""
      +"""(go as CompartmentHotSpotDefinition).aSet.size ==
          1"""
      +"""(go as CompartmentHotSpotDefinition).index.indexUL
          != (0 to 0)"""
    }

    def("DashedRegion") {
      +"""go is CompartmentHotSpotDefinition"""
      +"""
        go as CompartmentHotSpotDefinition
        go.getASet()
            .any {
              it hasPattern "BaseRegion" ||
              it hasPattern "DashedRegion"
            }
      """
      +"""(go as CompartmentHotSpotDefinition).isDashed()"""
    }
```

```

  def("Region") {
    +"""go hasPattern "BaseRegion" || go hasPattern "DashedRegion" """
    +"""go.allFromRelations(CompartmentHotSpotDefinition::class.java).isEmpty()"""
  }

  def("Transition") {
    +"""go.shapeClass == StraightSegmentLine::class.java"""
    +"""go.relatedTo(ConnectedBy::class.java) is ConnectionTo"""
  }
}

displayManagers {
  def("State", "State") {
    +"""
      val hsds = dm.go.allFromRelations(CompartmentHotSpotDefinition::class.java)
      dm.me.name = when {
          dm.go.hasPattern("SimpleState") -> dm.go.relatedTo(LabelFor::class.java)
          hsds.isNotEmpty() -> hsds.find { it.index.indexUL == (0 to 0) }!!
                                     .relatedTo(LabelFor::class.java)
          else -> dm.go
                      .allRelatedFrom(Contains::class.java)
                      .filter { it.shape is Label }
                      .map { it as GraphicObject<Label> }
                      .firstOrNull {
                            it.allFromRelations(LabelFor::class.java)
                              .none { r -> r.probability!!.probability > 0.9 }
                      }
      }
          ?.shape
          ?.text
    """

    +"""
```

```
        dm.me.regions = dm.go
            .allFromRelations(CompartmentHotSpotDefinition::
               class.java)
            .filterPattern("Region")
            .mapNotNull { it.dm<Region>()?.me?.ref() }
      """

      +"""
        dm.me.elements = dm.go
            .allRelatedFrom(Contains::class.java)
            .filterPattern("StateElement")
            .mapNotNull { it.dm<StateElement>()?.me?.ref() }
      """
    }

    def("Region", "Region") {
      +"""
        dm.me.elements = dm.go
            .allRelatedFrom(Contains::class.java)
            .filterPattern("StateElement")
            .mapNotNull { it.dm<StateElement>()?.me?.ref() }
      """
    }

    def("Transition", "Transition") {
      +"""
        dm.me.event = dm.go
            .allToRelations(LabelFor::class.java)
            .filter { it.probability!!.probability > 0.9 }
            .map { it.getA() }
            .filter { haveSameParent(it, dm.go) }
            .takeIf { it.size == 1 }
            ?.first()
            ?.shape
            ?.text
      """

      +"""
        val connection = dm.go.relatedTo(ConnectedBy::class.
           java) as? ConnectionTo
        val getSE = {go: GraphicObject<*>? ->
          go?.takeIf { it.hasPattern("StateElement") }
            ?.dm<StateElement>()?.me?.ref()
```

```
        }
        dm.me.source = getSE(connection?.getA())
        dm.me.target = getSE(connection?.getB())
      """
    }
  }
}
```

Listing C.6: `ConcreteMM` definition for Statecharts

D Test Setup

The test setup in this section is an exact replica of the StateGridTest described in [1]. This test serves as common ground to compare COUCHEDIT's reference implementation if the implementation produced in this Thesis.

D.1 StateGridTest

D.1.1 Scenario Parameters

- `gridSizeX` and `gridSizeY`: Number of rows and columns in the generated state representation grid, respectively.
- `numberOfStatesToMove`: Amount of state representations in the concrete syntax that will be selected in step 3 and moved to the right of all other state representations.
- `additionalGridSizeX and additionalGridSizeY`: Size of the grid of additional states that will be inserted in step 5.
- `toRemoveGridSizeX` and `ToRemoveGridSizeY`: Number of columns respectively rows of the grid that should be deleted in its top left corner in step 7.

D.1.2 Test Steps

1. **Process**: Generate a grid of `gridSizeX` x `gridSizeY` concrete syntax state representations (rounded rectangle + label) and insert them as one DiffCollection.
2. **Assert** that for every concrete syntax representation, a State abstract syntax Element has been created and that the State name property is equal to the concrete syntax representation's label's content text.
3. **Process**: Move a set of `numberOfStatesToMove` concrete syntax state representations out of the middle of the grid so that they now are to the right of the previously right border of the state representation grid.
4. **Assert** that all moved Elements still represent the same state, as such purely graphical changes should not change the abstract syntax representation of the statechart.

5. **Process**: Insert a new grid of `additionalGridSizeX` x `additionalGridSizeY` state representations to the bottom right of the existing grid.
6. **Assert** that for all of the new grid elements of step 5, an abstract syntax State Element has also been inserted.
7. **Process**: Remove `toRemoveGridSizeX` x `toRemoveGridSizeY` state representations from the top left of the grid.
8. **Assert** that for all of the removed state representations, the corresponding State abstract syntax Element has also been deleted.

List of Figures

List of Tables

List of Listings

Bibliography

[1] L. Nachreiner, “CouchEdit - a modular graphical editing architecture for flexible modeling,” Masters Thesis, University of Ulm, Ulm, Feb. 17, 2020. [Online]. Available: http://dx.doi.org/10.18725/OPARU-25291.

[2] Y. Van Tendeloo, S. Van Mierlo, B. Meyers, and H. Vangheluwe, “Concrete syntax: A multi-paradigm modelling approach,” in *Proceedings of the 10th ACM SIGPLAN International Conference on Software Language Engineering - SLE 2017*, Vancouver, BC, Canada: ACM Press, 2017, pp. 182–193, ISBN: 978-1-4503-5525-4. DOI: 10.1145/3136014.3136017. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3136014.3136017 (visited on 03/15/2020).

[3] V. Vaishnavi and W. Kuechler. (Jan. 2004). “Design research in information systems,” [Online]. Available: http://desrist.org/design-research-in-information-systems.

[4] D. Harel, “Statecharts: A visual formalism for complex systems,” *Science of computer programming*, vol. 8, no. 3, pp. 231–274, 1987, Publisher: Elsevier.

[5] D. Wüest, N. Seyff, and M. Glinz, “FlexiSketch team: Collaborative sketching and notation creation on the fly,” in *Proceedings of the 37th International Conference on Software Engineering - Volume 2*, ser. ICSE ’15, Florence, Italy: IEEE Press, May 16, 2015, pp. 685–688. (visited on 08/26/2020).

[6] G. Costagliola, A. Delucia, S. Orefice, and G. Polese, “A classification framework to support the design of visual languages,” *Journal of Visual Languages & Computing*, vol. 13, no. 6, pp. 573–600, Dec. 2002, ISSN: 1045926X. DOI: 10.1006/jvlc.2002.0234. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1045926X0290234X (visited on 04/27/2020).

[7] D. Moody, “The “physics” of notations: Toward a scientific basis for constructing visual notations in software engineering,” *IEEE Transactions on Software Engineering*, vol. 35, no. 6, pp. 756–779, Nov. 2009, ISSN: 0098-5589. DOI: 10.1109/TSE.2009.67. [Online]. Available: http://ieeexplore.ieee.org/document/5353439/ (visited on 04/23/2020).

[8] P. Bottoni and A. Grau, “A suite of metamodels as a basis for a classification of visual languages,” in *2004 IEEE Symposium on Visual Languages - Human Centric Computing*, Rome: IEEE, 2004, pp. 83–90, ISBN: 978-0-7803-8696-9. DOI: 10.1109/VLHCC.2004.5. [Online]. Available: http://ieeexplore.ieee.org/document/1372303/ (visited on 04/23/2020).

[9] A. Schürr, "Specification of graph translators with triple graph grammars," in *Graph-Theoretic Concepts in Computer Science*, E. W. Mayr, G. Schmidt, and G. Tinhofer, Eds., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 1994, pp. 151–163, ISBN: 978-3-540-49183-5. DOI: 10.1007/3-540-59071-4_45.

[10] A. Schürr and F. Klar, "15 years of triple graph grammars," in *Graph Transformations*, H. Ehrig, R. Heckel, G. Rozenberg, and G. Taentzer, Eds., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2008, pp. 411–425, ISBN: 978-3-540-87405-8. DOI: 10.1007/978-3-540-87405-8_28.

[11] F. Fondement and T. Baar, "Making metamodels aware of concrete syntax," in *Model Driven Architecture – Foundations and Applications*, vol. 3748, Series Title: Lecture Notes in Computer Science, Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 190–204, ISBN: 978-3-540-30026-7 978-3-540-32093-7. DOI: 10.1007/11581741_15. [Online]. Available: http://link.springer.com/10.1007/11581741_15 (visited on 04/23/2020).

[12] T. Baar, "Correctly defined concrete syntax," *Software & Systems Modeling*, vol. 7, no. 4, pp. 383–398, Oct. 2008, ISSN: 1619-1366, 1619-1374. DOI: 10.1007/s10270-008-0086-z. [Online]. Available: http://link.springer.com/10.1007/s10270-008-0086-z (visited on 04/27/2020).

[13] M. Minas and G. Viehstaedt, "DiaGen: A generator for diagram editors providing direct manipulation and execution of diagrams," in *Proceedings of Symposium on Visual Languages*, ISSN: 1049-2615, Sep. 1995, pp. 203–210. DOI: 10.1109/VL.1995.520810.

[14] M. Minas, "Concepts and realization of a diagram editor generator based on hypergraph transformation," *Science of Computer Programming*, vol. 44, no. 2, pp. 157–180, Aug. 2002, ISSN: 01676423. DOI: 10.1016/S0167-6423(02)00037-0. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167642302000370 (visited on 04/07/2020).

Name: Florian Lappe Matrikelnummer: 922114

Erklärung

Ich erkläre, dass ich die Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Ulm, den ....

Florian Lappe