



**HOCHSCHULE KONSTANZ TECHNIK, WIRTSCHAFT UND GESTALTUNG**  
UNIVERSITY OF APPLIED SCIENCES

# Scala als Generatorsprache

**Julian Müller**

**Konstanz, 15.02.2016**

## BACHELORARBEIT

# BACHELORARBEIT

zur Erlangung des akademischen Grades

**Bachelor of Science (B. Sc.)**

an der

**Hochschule Konstanz**

Technik, Wirtschaft und Gestaltung

**Fakultät Informatik**

Studiengang Angewandte Informatik/Softwareengineering

Thema: **Scala als Generatorsprache**

Bachelorkandidat: Julian Müller, Mondrauteweg, 78467Konstanz

1. Prüfer: Prof. Doc. Marko Boger

2. Prüfer: Titel, Markus Gerhart

Ausgabedatum: [Datum]

Abgabedatum: 15.02.2016

## Zusammenfassung (Abstract)

Thema: Scala als Generatorensprache

Bachelorkandidat: Julian Müller

Firma: HTWG Konstanz

Betreuer: Prof. Doc. Marko Boger  
Titel, Markus Gerhart

Abgabedatum: 15.02.2016

Schlagworte: Scala, MoDiGen, Model-Driven Software Development (MDSD), Domain-Specific Language (DSL), Parser, Generatoren

Das MoDiGen Projekt will das Erzeugen graphischer Editoren schnell und unkompliziert ermöglichen. Hierfür werden nach einem Set textueller DSLs Modellklassen erzeugt, welche die graphischen Elemente abbilden. Bisher wurde Xtext benutzt um die dafür nötigen Objektstrukturen, Parser und Generatoren erzeugen zu lassen. Da Das Projekt langsam kontinuierlich auf Scala als Programmiersprache umsteigt, soll nun untersucht werden, ob die gegebenen Funktionalitäten von Xtext in Scala umgesetzt werden können und ob Features implementiert werden können, die Xtext nicht bietet. Es werden also die gewünschten Objektstrukturen der Modellklassen umgesetzt. Außerdem werden die nötigen Parser und Generatorenklassen erstellt. Dabei spielen Vorgaben wie Vererbung und die transitive Weitergabe der Objekteigenschaften eine wichtige Rolle.

# Ehrenwörtliche Erklärung

Hiermit erkläre ich *Julian Müller, geboren am 07.07.1991 in Donaueschingen*, dass ich

- (1) meine Bachelorarbeit mit dem Titel

## **Scala als Generatorensprache**

bei der HTWG Konstanz unter Anleitung von Prof. Doc. Marko Boger selbständig und ohne fremde Hilfe angefertigt und keine anderen als die angeführten Hilfen benutzt habe;

- (2) die Übernahme wörtlicher Zitate, von Tabellen, Zeichnungen, Bildern und Programmen aus der Literatur oder anderen Quellen (Internet) sowie die Verwendung der Gedanken anderer Autoren an den entsprechenden Stellen innerhalb der Arbeit gekennzeichnet habe.

Ich bin mir bewusst, dass eine falsche Erklärung rechtliche Folgen haben wird.

Konstanz, 15.02.2016

---

(Unterschrift)

# Inhaltsverzeichnis

<b>Ehrenwörtliche Erklärung</b>	<b>I</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Was Ist MoDiGen? . . . . .	2
1.2 Aufgabenstellung . . . . .	3
<b>2 Grundlagen: Scala/Xtext</b>	<b>5</b>
2.1 Scala . . . . .	5
2.2 Xtext . . . . .	6
<b>3 Ansatz</b>	<b>7</b>
3.1 Objektstruktur . . . . .	7
3.1.1 ClassHierarchy . . . . .	7
3.2 Parser . . . . .	8
3.2.1 Parsing Mechanismus und Parserlogik . . . . .	10
3.2.2 Factory Klassen . . . . .	10
3.2.3 Parsing Ablauf . . . . .	11
3.3 Auflösung von Vererbungshierarchien . . . . .	13
3.3.1 Kennzeichnung nicht definierter Felder in Modellklassen .	14
3.3.2 Vererbungsmechanismus über Referenz auf Elterninstanzen	15
3.3.3 Vererbungsmechanismus durch direkte Weitergabe der At- tribute bei Instantiierung . . . . .	15
3.4 Generatoren . . . . .	16
<b>4 Umsetzung</b>	<b>17</b>
4.1 Objektstruktur . . . . .	17
4.1.1 Modellklassen . . . . .	17
4.1.2 Besonderheiten bei Style . . . . .	17
4.1.3 Besonderheiten bei Shape . . . . .	20
4.1.4 Besonderheiten bei Diagram . . . . .	22
4.1.5 Vererbung/ClassHierarchy . . . . .	24
4.2 Implementierung der Parserklassen . . . . .	25
4.2.1 Factory Klassen . . . . .	25
4.2.2 Problematik der rekursiven regulären Ausdrücke . . . . .	26
4.2.3 Lösung Parser Combinators . . . . .	28
4.2.4 Prinzip der Parserlogik . . . . .	29
4.3 Vererbung . . . . .	32

4.3.1	Vererbung bei Shapes . . . . .	34
4.3.2	Transitivität der Styleeigenschaften . . . . .	35
4.3.3	Modell und Skizze; Problematik der Instantiierungsreihenfolge . . . . .	38
4.3.4	Generatorenklassen . . . . .	42
4.4	Ausführliches Beispiel . . . . .	43
4.5	Cache . . . . .	46
4.5.1	Implicits . . . . .	47
<b>5</b>	<b>Fazit . . . . .</b>	<b>49</b>
5.1	Probleme . . . . .	49
5.1.1	Scalaspezifische Probleme . . . . .	49
5.2	Schlussfolgerung/Ausblick . . . . .	51
5.2.1	Was wurde erreicht . . . . .	51
5.2.2	Was nicht erreicht wurde . . . . .	51
5.2.3	Ausblick . . . . .	52

# Kapitel 1

## Einleitung

Seitdem Rechner mit Maschinencode programmiert werden, steigt das Abstraktionslevel der Programmiersprachen kontinuierlich an (siehe Abbildung 1.1). Wer die Assemblersprachen kennt und fürchtet, weiß Programmiersprachen wie Fortran, Pascal oder C zu schätzen, da sie Kontrollstrukturen einführen, welche leicht verstanden und eingesetzt werden können. Wer stets wiederkehrende Muster programmiert, bedient sich der Objektorientierten Sprachen, wie zum Beispiel Java oder C++. Indem die Logik hinter Programmiersprachen den Paradigmen angeglichen werden, die die Menschen intuitiv (oder am einfachsten) beherrschen, wird Programmieren erheblich effizienter, sowohl im Lernprozess, als auch in der Umsetzung. Der prozentuale Anteil an wiederverwendbarem Code in einem Projekt steigt mit dem Abstraktionslevel der verwendeten Programmierparadigmen. Code, der wiederverwendet werden kann, spart Zeit und reduziert die Chance neue Fehler in das Projekt einzuarbeiten. Die Objektorientierung wurde eingeführt um wiederkehrende Elemente einheitlich behandeln zu können (z.B. graphische Elemente), somit die Komplexität des Codes zu verringern und dessen Wartbarkeit zu erleichtern. Die Programmiersprache *Scala* ist ebenfalls objektorientiert, erfüllt jedoch ebenso die Kriterien einer funktionalen Programmiersprache und bietet zudem einige Features, die sie besonders dafür qualifizieren, sogenannte *Domain Specific Languages* zu erstellen. Dies begünstigt die Unterstützung des nächsten Schrittes in der Evolution der Programmierparadigmen. Nicht zuletzt der Markt bekräftigt den Wunsch nach eben dieser Effizienz und fordert diese nächste Abstraktionsstufe. Entwickler auf der ganzen Welt benutzen Bibliotheken und Frameworks, um bereits entwickelte Lösungen wiederverwenden zu können und um bekannte Lösungsansätze für bekannte Problemstellungen zu verwenden. Der hierbei entwickelte Code richtet sich also nach einem bereits vorgelegten Plan. Jener Quellcode, der manuell ergänzt werden muss, lässt sich nun teilweise automatisiert ergänzen. Dieses Ziel verfolgt das **Model Driven Software Development** (*MDSD*). Schematischer Code kann in domänenspezifischen Modellen verarbeitet werden. Nachdem Modelle konzipiert sind, ist es möglich automatisiert auf domänenspezifische Problemstellungen einzugehen. Im Falle der Industrie bedeutet dies Massenproduktion, ermöglicht von Robotern. In Bezug auf Softwareentwicklung, bedeutet dies Code, der Code generiert. Die Idee ist, va-

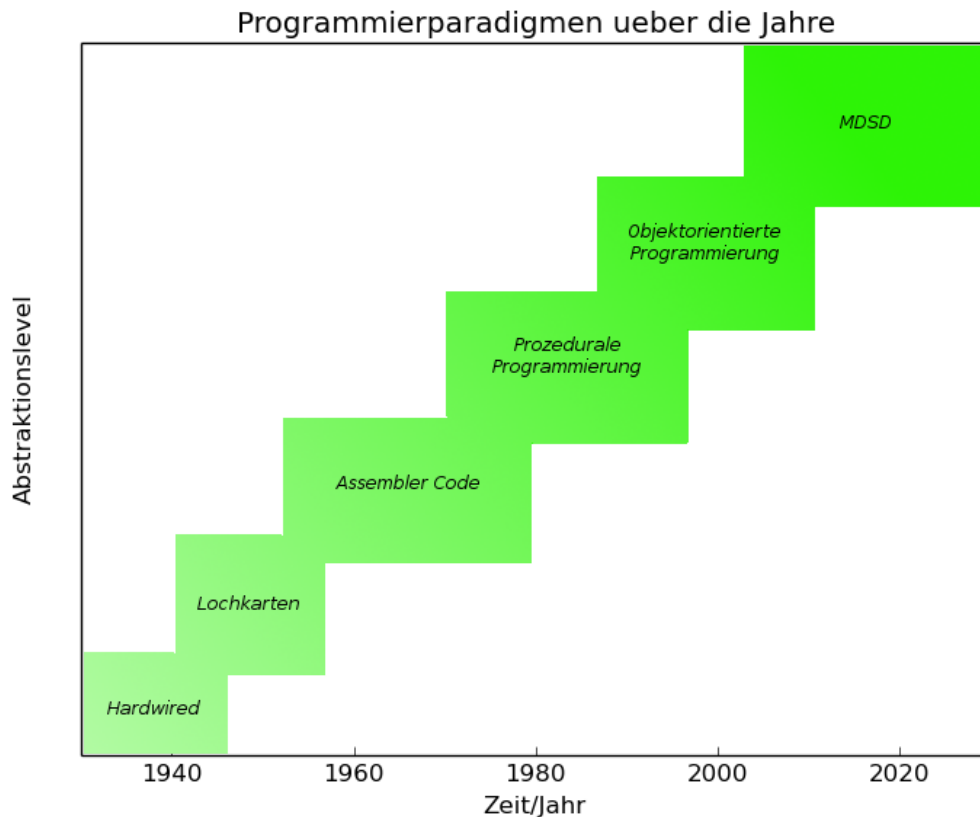


Abbildung 1.1: Programmierparadigmen im Verlauf der Zeit

riable Teile des Codes über ein Modell zu beschreiben. Das Programm ist selbstständig in der Lage, das Modell in den gewünschten Code umzuwandeln. Das besagte Modell kann hierbei drastisch vereinfacht, oder auf eine bestimmte Benutzerzielgruppe(Domäne) angepasst sein. Es lässt sich also problemlos einrichten, beispielsweise einem Banker über seine gewohnten Fachbegriffe eine **Domain Specific Language (DSL)** zu bieten, welche syntaktisch auch noch einer seiner Tabellen oder beispielsweise Buchungssätzen ähnelt und somit für ihn unkompliziert und sofort erlernbar ist. Für die Umwandlung eines Modells in konkreten Code (Generat) sind sogenannte Generatoren notwendig. Siehe auch Trompeter and Beltran 2007

## 1.1 Was Ist MoDiGen?

Domänenspezifische Modellierung erfreut sich wachsender Beliebtheit in der Softwareentwicklung. Projekte, die für ein domänenspezifisches Problem graphische Modellierungstools benötigen, können entweder ein sehr spezifisches Tool benutzen, was nur dann möglich ist, wenn das Problem exakt durch das Tool beschrieben werden kann. Ansonsten muss ein generischer Editor herange-



zogen werden, doch diese sind dann in der Regel derart komplex, dass es nicht mehr möglich ist, sein Problem schnell und unkompliziert darstellen zu können. Das *MoDiGen* (**M**odel **D**iagram **G**enerator) Projekt will es ermöglichen, durch kurze und schnell umsetzbare textuelle DSLs, graphische Editoren erzeugen zu lassen, deren Entwicklung sonst überaus zeitaufwendig sein kann. Hierfür wurde einerseits ein hoch abstrahiertes Metamodell erstellt, das komplexe Klassenhierarchien darstellen kann, außerdem wurden einige textuelle DSLs erstellt, über die die Modellklassen beschrieben werden können (Diagram, Shape, Connection, Style). So soll dem Anwender (*Domain Expert*) die Möglichkeit geboten werden, eigene graphische Elemente zu erzeugen. Dabei kann es sich auch um Verbindungselemente handeln, denn diese werden im Metamodell von MoDiGen getrennt und ebenbürtig zu den eigentlichen Elementen gehandhabt (vgl. Gerhart and Boger (**year?**)). Außerdem werden sie nicht wie bei etwa *Ecore* nur als Attribut eines Modells gehalten. Im Hinblick auf die Skalierbarkeit ist dies vorteilhaft, da Elemente nicht über ein relationales Schema abgebildet werden müssen, um sie in einer Datenbank zu hinterlegen. Hieraus entstehen wiederum Vorteile bezüglich der Skalierbarkeit und des Datenzugriffs.

## 1.2 Aufgabenstellung

Hinsichtlich der Flexibilität der Modelle und der Skalierbarkeit, bieten bestehende Lösungen wie *Xtext* und *Ecore* derzeit keine Ideale Lösung. Die relativ junge Programmiersprache *Scala* ist, wie der Name schon andeuten soll, mit einem besonderen Fokus auf die Skalierbarkeit entwickelt worden. Wie bereits in 1.1 erläutert, stützt sich das MoDiGen Projekt auf ein neues, selbst entwickeltes Metamodell, welches hinsichtlich der Skalierbarkeit und der Speicherauslastung wesentlich bessere Ergebnisse erzielt. So wie das Metamodell sollen nun auch die Parser und Generatoren in *Scala* umgesetzt werden. Es soll untersucht werden, inwiefern *Scala* (als allzweckstaugliche Programmiersprache) sich dafür eignet, die Rolle bestehender spezialisierter Technologien zu übernehmen. Wo kann *Scala* seine Stärken spielen lassen und wo ist es den bisherigen Lösungen unterlegen? Dafür sollen u.a. MoDiGens Modellklassen: Style, Shape und Diagram in *Scala* umgesetzt werden und die entsprechenden Parser und Generatoren erstellt werden. Konkreter: Es müssen zunächst die geforderten Objektstrukturen geschaffen werden um Style, Shape, Connection und Diagram abbilden zu können. Hierbei muss darauf geachtet werden, dass ein Style ein weiteres Style assoziieren kann um eine Vererbung zu ermöglichen.

Shapes können derzeit aufgrund der Komplexität der Modellklasse nicht vererbt werden. Sowohl die Shapevererbung als auch Mehrfachvererbung im allgemeinen ist mit bestehenden Technologien nicht möglich und soll daher auf Umsetzbarkeit mit *Scala* überprüft werden. Des weiteren sind die Beziehungen der Modelle untereinander genau geklärt - siehe Abbildung 1.2. Diagrams können auf Styles und Shapes verweisen und Shapes können Styles referenzieren. Außerdem muss ein Parser erstellt werden, der in der Lage ist die *Spray* DSL zu interpretieren und entsprechend der Modelle zu instantiieren. Zuletzt kommen

Generatorenklassen, die schließlich entsprechende Generate aus den erzeugten Objekten erstellen können.

Es handelt sich also um eine Studie zur Umsetzbarkeit von Parser- und Generatorenklassen in Scala, mit speziellem Fokus auf die Möglichkeit komplexe Vererbungshierarchien darstellen zu können. Hierbei liegt die größte Herausforderung darin, einen Parsingmechanismus zu erstellen, der mit rekursiven DSL Definitionen umgehen kann und der unabhängig davon um welche Modellklasse es sich handelt, in der Lage ist Vererbungshierarchien zu erkennen und umzusetzen. Da der Autor Scala erst für dieses Projekt erlernt, kann davon ausgegangen werden, dass alle verwendeten Techniken anfängerfreundlich sind.

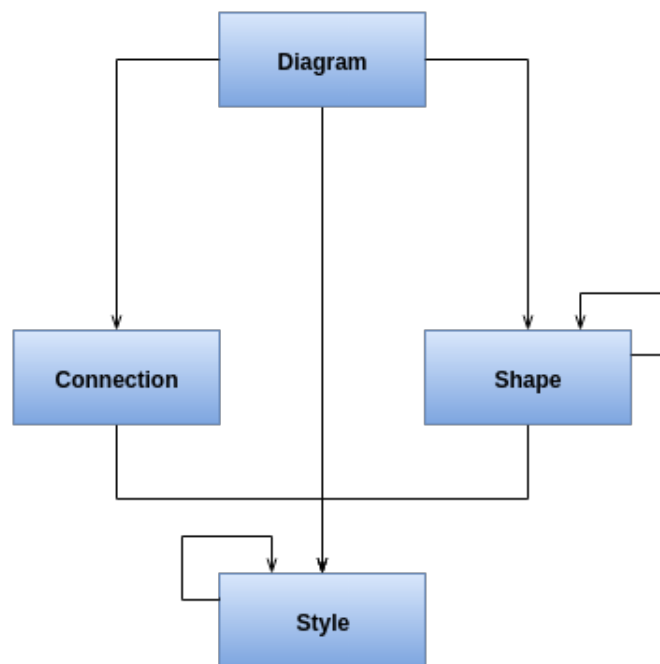


Abbildung 1.2: Klassendiagramm Diagram, Shape, Style (Sehr stark vereinfacht)

## Kapitel 2

# Grundlagen: Scala/Xtext

### 2.1 Scala

Im folgenden soll zum Verständnis zunächst die Programmiersprache *Scala* vorgestellt werden.

Die Entwicklung von *Scala* (Scalable Language) begann bereits 2001 an der *École polytechnique fédérale de Lausanne* in der Schweiz durch ein Team um Professor *Martin Odersky* (vgl. Braun 2011, S.1). Designziel war eine elegante, typsichere Programmierung, vollkommene Kompatibilität mit der *JVM* und eine Zusammenführung der Objektorientierung und der funktionalen Programmierung (vgl. Braun 2011, S.2). Scala ist hierbei objektorientierter als z.B. Java, da in Scala *alles* ein Objekt ist. Odersky selber bezeichnet Scala als postfunktionale Sprache (vgl. Braun 2011, S.2), da immer noch diskutiert wird, ob Scala als funktionale Programmiersprache betitelt werden darf. Fest steht jedoch, dass man in Scala funktional programmieren kann - man muss es allerdings nicht. Funktionen können Argumente oder Ergebnisse anderer Funktionen sein und sind somit *Higher Order Functions*. Trotz der Typsicherheit bietet Scala das *Feeling* einer dynamisch typisierten Sprache, da Typen beim Kompilieren inferiert werden können. Der *Compiler* erkennt also anhand des initialisierten Wertes, um was für einen Typ es sich handelt, so kann fast immer auf manuelle Typkennzeichnung verzichtet werden.

```
var name = "Julian"  
var alter = 24  
var nachname:String = "Müller"  
var matrikelnummer:Int = 287501
```

Listing 2.1: Beispielhafte Variablendefinition mit Typinferierung und manueller Typangabe

Die Variablendefinitionen *name* und *alter* aus Listing 2.1 sind vollkommen korrekte Ausdrücke und werden ebenso kompiliert wie *nachname* und *matrikelnummer*.

Scala ist einfacher zu erlernen, als beispielsweise Java, da Anweisungen auch alleinstehend als Skript oder direkt in einem interaktiven Scala Interpreter ausgeführt und ausprobiert werden können. Komplexe Problemstellungen können stark abstrahiert in wenigen Zeilen Code beschrieben werden, da Funktionen Objekte sind und als Parameter übergeben werden können. Über entsprechende Methoden, kann man sogar eigene Operatoren und Kontrollstrukturen erstellen. Übrigens: „*Damit ist Scala prädestiniert zur Erstellung von Domain Specific Languages [...]*“ (Braun 2011, S.2).

## 2.2 Xtext

Nachdem nun Scala erörtert wurde, soll im folgenden die zu ersetzende Technologie *Xtext* beschrieben werden.

Xtext ist ein Framework mit dem u.a. domänenspezifische Sprachen entwickelt werden können und ergänzt das *Eclipse-Modeling-Framework* (vgl. Bettini 2013, S.13). Dafür wird eine komplexe Grammatiksprache zur Verfügung gestellt, über die Entitäten deklariert und assoziiert werden können. Anhand der Grammatik erzeugt Xtext Parser, Generatoren und sogar ein Klassenmodell für die beschriebenen Klassen. Um die beschriebene DSL ausführbar zu machen wird außerdem ein Texteditor generiert, welcher entsprechendes *Syntax Highlighting* automatisch programmiert. Der große Vorteil von Xtext ist, dass die Definition der beschriebenen Sprache gleichzeitig die Objektstruktur der zu generierenden Modelle beschreibt. Wird die DSL verändert, ändert sich die Objektstruktur der Modelle ebenfalls und so bleibt die Abbildung von DSL auf Modell immer konsistent.

# Kapitel 3

## Ansatz

Nachdem die Problemstellung geklärt ist und die konkurrierenden Technologien erläutert wurden, sollen nun die zugrundeliegenden Gedanken zur Problemlösung vorgestellt werden. Die hier aufgeführten Konzepte für Beispiele verwendete Methoden- oder Klassennamen entsprechen nicht zwingend der tatsächlichen Umsetzung.

### 3.1 Objektstruktur

Zunächst werden die entsprechenden Modellklassen aus der Xtext Grammatik in Scala Quellcode übernommen. Style-, Shape-, Connection- und Diagram Klassen mit entsprechenden Feldern werden zunächst als sogenannte *case classes* erstellt. Dies hat den Vorteil, dass im Falle einer später notwendigen Verifizierung der Instanzen, vorimplementierte *hash-* *equals-* und *toString* Methoden zur Verfügung zu haben. Da Style am flexibelsten ist (nur eingehende Abhängigkeiten), wird Style.scala zuerst erstellt und entsprechend der Assoziationsreihenfolge dann Shape.scala und Diagram.scala (siehe Abbildung 1.2).

#### 3.1.1 ClassHierarchy

Da es für die Instanzen mancher Modellklassen möglich sein soll, Objekte der selben Klasse zu erweitern, ist die Idee des Autors, eine Collection namens *ClassHierarchy* zu erstellen, welche in der Lage ist, ähnlich eines Baumes Vererbungshierarchien in Form von Knoten (*Nodes*) darzustellen. Siehe dazu Abbildung 3.1. So werden die Modellklassen, für die die Vererbung möglich sein soll, über diesen Baum in Relation gesetzt. Im Falle einer gänzlich neuen Modellklasse (z.B. ein neuer Style), wird die Instanz als neue Basisklasse eingefügt. Ansonsten über eine Methode, die beispielsweise *inheritsFrom* heißen könnte. *InheritsFrom* muss entweder von der *ClassHierarchy* oder den Knoten selber zur Verfügung gestellt werden. Über die entsprechende Argumentenliste wird vermittelt, welche Knoten erweitert werden. *InheritsFrom* ergänzt dann eigenständig die *parent-* und *children* Listen eines Knotens um die neue Instanz der Modellklasse.

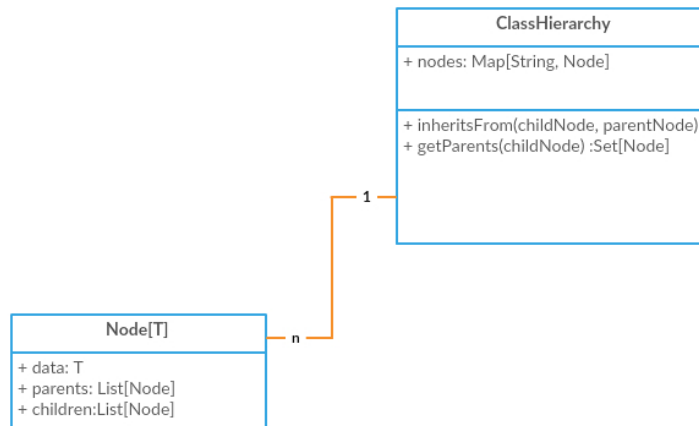


Abbildung 3.1: Klassenmodell von ClassHierarchy

Wie in 3.1 zu sehen ist verfügen die Nodes über Listen, welche auf Eltern- und auf Kind knoten verweisen. So kann man bequem von jedem Style erfahren, welche Styles er erweitert und welche Styles ihn erweitern. Außerdem enthält *ClassHierarchy* eine Map, die Namen (Strings) auf beliebige Klassen abbilden kann. Über dieses mapping ist ein schneller Zugriff auf die gewünschte Klasse möglich, ohne mühsam durch eine Baumstruktur der Nodes navigieren zu müssen. Um dieses Prinzip auf alle möglichen Klassen anwenden zu können, sollte *ClassHierarchy* generisch sein. Vorausgehend waren die Anforderungen für Vererbung nur auf Style bezogen. Da die Vererbungsprinzipien aber auch bei Shape hilfreich sein könnten, ist so gewährleistet, dass man sich diese Tür nicht unnötig verschließt. Benötigte Enumerationen und andere nicht komplexe Hilfsobjekte der Klassen werden in dessen *Companion Objekt* definiert. Über entsprechende *Apply*-Methoden wird der Aufruf der *ClassHierarchy* so einfach wie möglich gestaltet (siehe Listing 3.1, Zeile 2).

```

classHierarchy.nodes.get("styleName").data
classHierarchy("styleName")

```

Listing 3.1: Beispielhafte Vereinfachung durch Apply Methode

Zeile 1 in Listing 3.1 zeigt wie der Aufruf eigentlich ausssehen müsste. Zeile 2 hingegen ist kürzer und außerdem verständlicher. Da der Aufruf in diesem Fall sehr eindeutig ist, kann ruhigen Gewissens auf die Apply-Methode zurückgegriffen werden. „Fazit: Die Methoden *apply* und *update* sollten nur eingesetzt werden, wenn intuitiv klar ist was der Einsatz der Klammern bedeutet“ (Esser 2011, S.76).

## 3.2 Parser

Neben den Gedanken zur Vererbung der Modellklassen, ist es auch wichtig zu klären, wie die Modelle überhaupt aus der Stringform in Objektform umzuwan-

deln sind. Entsprechende Mechanismen werden allgemein als Parser bezeichnet. Dieser Parser ist verantwortlich dafür die in Stringform enthaltenen Styles, Shapes, Connections und Diagrams als diese zu identifizieren und in Objekte umzuwandeln. Die Struktur eines Styles (ähnlich auch bei den anderen Modellen) ist in Listing 3.2 beispielhaft dargestellt.

```
style BpmnDefaultStyle {  
  
    description = "The default style of the petrinet  
                  diagram type."  
  
    transparency = 0.95  
  
    background-color = black  
  
    line-color = black  
  
    line-width = 1  
  
    font-color = black  
  
    font-name = "Tahoma"  
  
    font-size = 6  
  
    font-bold = yes  
  
    font-italic = yeshttp://www.macwrench.de/wiki/  
                  Kurztipp\_-\_Quellcodelistings\_in\_LaTeX  
  
    gradient-orientation = horizontal  
    und  
}
```

Listing 3.2: Beispielhafte Style Definition über die entsprechende Style DSL

Die Xtext Grammatik soll außerdem um eine Vererbungskomponente erweitert werden, die einerseits auch Shapes die Möglichkeit bietet andere Shapes zu erweitern. Zusätzlich soll auch Mehrfachvererbung untersucht und gegebenenfalls umgesetzt werden. Shapes sollen also so definiert werden können wie in Listing 3.3 zu sehen ist.

```
shape SomeShape extends StandardShape , AnotherShape { ... }
```

Listing 3.3: (Mehrfach)Vererbung bei einer Shape Definition

Es ist zwar bereits möglich, Style zu erweitern, doch was die Vererbung von Shapes und Mehrfachvererbung generell angeht besteht noch Bedarf, da Xtext nicht in der Lage ist, diese Features umzusetzen.

### 3.2.1 Parsing Mechanismus und Parserlogik

Im Folgenden wird eine konkrete Technik vorgestellt, mit der das Umwandeln von Strings in Objekte erfolgen kann, die sogenannten *Regulären Ausdrücke*. Über diese regulären Ausdrücke kann der Style gefiltert werden, um so an der Syntax der DSL vorbei zu kommen und an die eigentlichen Attribute zu gelangen. Anhand des ersten Wortes, soll der Parser erkennen, um welchen Modelltyp (Style, Shape, Diagram) es sich handelt. Weiter würde ein vordefiniertes Schlüsselwort *extends* eine zu erweiternde Instanz kennzeichnen. Im Gegensatz zu diesen hart kodierten Schlüsselwörtern, werden eigentliche Attribute wie zum Beispiel *line-width* zunächst unter einheitlichen Regeln geparkt und erst in einer tieferen Parser klasse oder dem entsprechenden Konstruktor aufgelöst. Auf diese Art und Weise reichen wenige Regeln aus, um komplexe Ausdrücke abbilden zu können. So geparkte Attribute bleiben zunächst in der Stringform. Ziel dieses Vorgehens ist es zunächst nur Attributsname und Attributswert zu ermitteln, sicherzustellen, dass Syntaxregeln eingehalten wurden und die Weitergabe eines einheitlichen Datentyps an tiefer liegende Parserklassen. Diese erste Schicht des Parsers wird also noch keine Objekte, oder gar Baumstrukturen erzeugen, sondern lediglich alle Informationen aus dem String extrahieren, die an einen spezifischen und spezialisierten Parser weiter gegeben werden kann. Logisch wird hierdurch das Prinzip sogenannter *Lexer* dargestellt, welche insbesondere bei *Compilern* Anwendung finden (vgl. Aho 2008, S.135). Im Prinzip wird aber nichts anderes gemacht als eine vordefinierte Sprache zu lesen und umzusetzen, weshalb das Prinzip des Lexers auch hier sinnvoll ist. Die vordefinierte DSL kann so auch im Bezug auf die Modellklassen leichter modifiziert werden, da solange die Syntaxregeln der DSL nicht geändert werden, der Lexer von neuen Attributen unbeeindruckt ist. Erst die spezifischen Parser müssten bei einer Solchen Änderung um eine Regel erweitert werden, welche in der Lage ist die eigentliche Information aus dem neuen Argument zu extrahieren. Die wenig komplexen Styles sollten so einfach zu parsen sein. Shapes stellen den schwierigsten Teil beim Parsen dar, da geometrische Figuren(*GeometricModel*), wie Ellipsen oder Polygone ineinander geschachtelt werden können.

### 3.2.2 Factory Klassen

Es wird davon ausgegangen, dass Modellklassen durch Umwandlung aus einem bereitgestellten Stringinput erzeugt werden. Dieser Stringinput enthält alle notwendigen Informationen um eine einsatzfähige Instanz einer Modellklasse erzeugen zu können. Es ist wünschenswert die Erzeugung der Instanzen so konsistent wie möglich zu halten. Dies bedeutet möglichst wenige und klar definierte Wege ein Objekt zu erzeugen. Dies kann effektiv über eine Fabrik Klasse (zu Englisch *Factory Class*) erreicht werden.



Zur Erzeugung der Exemplare der Modellklassen wird so, wenn möglich, nur eine einzige Methode definiert, die schlussendlich auf den Konstruktor der Modellklasse verweist. Dies schließt allerdings noch nicht aus, dass Modellklassen auch anders erzeugt werden können. Scala stellt einen von Haus aus eingebauten Mechanismus zur Verfügung, die *Companion Objekte*. Scala hat sich von den in Java typischen statischen Variablen (Klassenvariablen) getrennt. Um das Prinzip der Objektorientierung konsequenter durchzusetzen, stellt es hierfür die Definition von *Singleton Objekten* zur Verfügung. Über das Schlüsselwort *object* kann anstatt einer Klassen-, eine Objektdefinition erfolgen. Jegliche Felder eines solchen Objekts sind (in Java Manier ausgedrückt) Instanzvariablen, jedoch importierbar. Allein die Möglichkeit auf diese Art und Weise *Singleton Objekte* zu erstellen, deckt bereits eine Funktionalität der Fabrik Klassen ab. Wird ein solches Objekt in der selben Quelldatei wie eine Klassendefinition erstellt und haben Klasse und Objekt außerdem den gleichen Namen, handelt es sich um *Companion Objekt* und *Companion Klasse*. Auf den ersten Blick, sind *Companion Objekte* nur der Ansatz Klassenvariablen bereitzustellen und trotzdem dem objektorientierten Paradigma treu zu bleiben. Ein entscheidender Vorteil ist jedoch, dass so definierte Objekte auch auf die privaten Member der *Companion Class* zugriff haben. Auf unseren Fall bezogen bietet sich also die Möglichkeit, den Konstruktor der Modellklasse mit dem Schlüsselwort *private* zu kennzeichnen. Folglich kann nur noch das *Companion Objekt* darauf zugreifen und man kann exklusiven Zugriff über eine Fabrik Methode des Companion Objekts anbieten.

### 3.2.3 Parsing Ablauf

Da in den bisherigen Gedanken geklärt wurde, wie Strings eingelesen werden, die Vererbung unter den Modellklassen dargestellt wird und wie Modellklassen konkret erzeugt werden, kann nun anhand des Sequenzdiagramms 3.2 beschrieben werden, wie ein Parsing Vorgang prinzipiell abzufließen hat. Erhält die Applikation einen String um beispielsweise eine Shape einzulesen, werden zunächst sämtliche Attribute aus dem String gefiltert. Hierzu zählen auch Name und erweiterte Shapereferenzen. Dies wird über eine Methode namens *parseArguments* (siehe Abbildung 3.2) erfolgen. Egal um welche Modellklasse es sich handelt, werden alle ausgelesenen Attribute in Stringform weiterverarbeitet. So wird die Verantwortung der Auflösung auf die einzelnen Parser der Modellklassen ausgelagert und diesen gleichzeitig ein einheitliches Medium geliefert. Auf diese Weise können ähnliche oder gleiche Attribute in unterschiedlichen Parsern, auf die selbe Weise aufgelöst werden. Wie bereits angedeutet, geht es von hier an mit einem auf die Modellklasse spezialisierten Parser weiter, der *ModelClassParser* (siehe Abbildung 3.2). Der spezifische Modellklassenparser muss nun zunächst prüfen, ob weitere Instanzen, des Zieltyps erweitert werden sollen. Hierfür gibt er die in den Attributen enthaltene Liste an Eltern (in Form von Strings) an die *ClassHierarchy* weiter, die den Überblick über die Vererbungshierarchie hat und die Namen über ein internes Mapping auf bestehende Referenzen auflösen kann (*resolveDependencyName*, vgl. 3.2). Sind alle Elternteile aufgelöst, kann der modellklassenspezifische Parser nun anhand einer

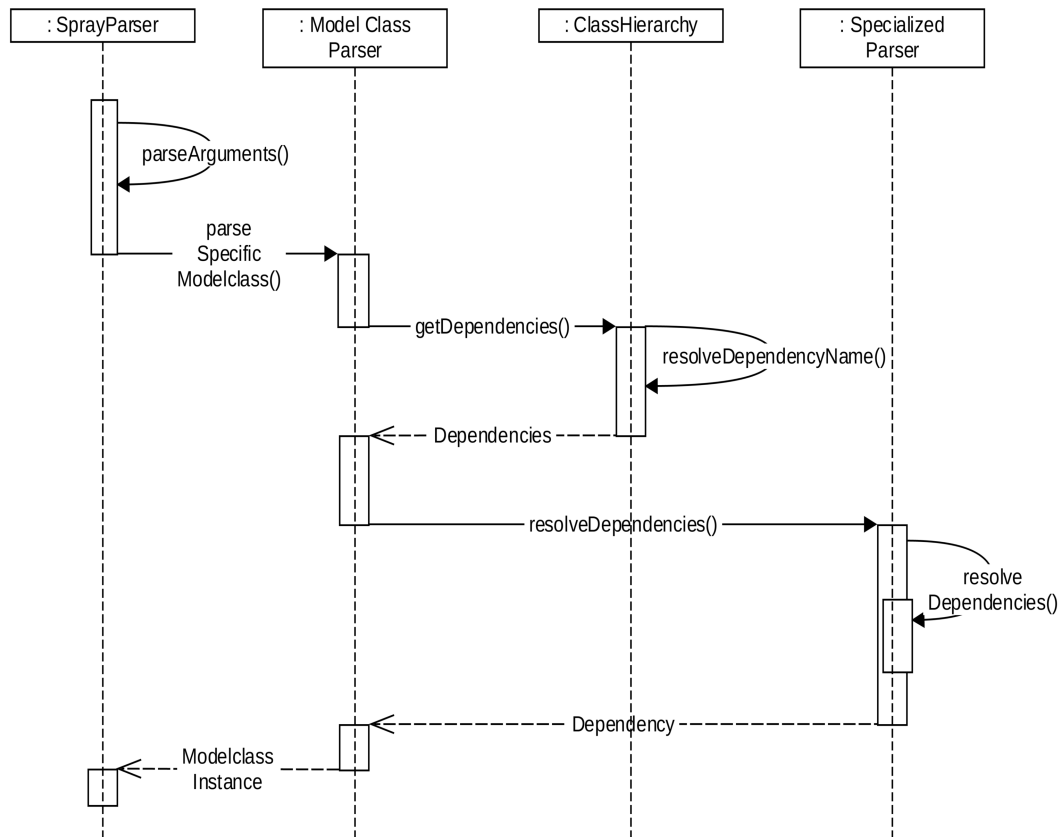


Abbildung 3.2: Sequenzdiagramm für den Ablauf eines Parsing Vorgangs

kompletten Liste an Attributen beginnen, die in Stringform enthaltenen Informationen der Abhängigkeiten aufzulösen. Falls es sich bei den in Stringform enthaltenen Informationen nicht um primitive Datentypen, sondern um komplexe eventuell ineinander geschachtelte Elemente handelt, werden diese wiederum an ihren speziell definierten Parser weitergegeben (*resolveDependencies*, vgl. 3.2). Da die Elemente wie bereits erwähnt eventuell ineinander verschachtelt sind, können viele weitere rekursive Aufrufe auf spezialisierte Parsereinheiten erfolgen, bis schließlich alle Argumente aufgelöst sind und eine fertige Instanz an den *SprayParser* zurück geliefert werden kann.

Die Parserlogik teilt sich also auf viele kleine spezialisierte Parser auf. Dabei gibt es einerseits den *SprayParser*, der die Schnittstelle zu allen vorhandenen Parsern darstellt. Außerdem die vier spezialisierten Style-, Shape-, Connection- und Diagrammparser, die die Modellklassen erzeugen und einige weitere kleinere Parser, welche dafür verantwortlich sind, eigens erstellte Datentypen parsen und erzeugen zu können, wie zum Beispiel *Anchors*.

### 3.3 Auflösung von Vererbungshierarchien

Ist eine Technik gefunden, mit der sich die DSL in die Objektstruktur der Modellklassen umwandeln lässt, muss direkt die Problematik der Vererbung betrachtet werden, da die hierfür benötigten Techniken schon beim Parsen relevant sind. Es soll ein Vererbungsmechanismus implementiert werden. Entgegen der geläufigen Definition der Vererbung, handelt es sich hierbei nicht um das Erweitern von Klassen, so wie man es aus diversen Programmiersprachen kennt. Die Anforderung definiert, dass einzelne Instanzen der Modellklassen in der Lage sein sollen, Objekte der selben Klasse erweitern zu können. Genauer bedeutet dies, dass eine Instanz **nicht** in der Lage sein soll dynamisch komplett neue Attribute zu definieren, sondern **bekannte** Attribute von Elterninstanzen zu übernehmen, beziehungsweise zu überschreiben. „Erweitern“ soll in diesem Kontext also zwei Situationen beschreiben:

- Elternteil verfügt über Attribute, die in der erweiternden Klasse nicht vorhanden sind ->Wert wird geerbt
- Elternteil verfügt über Attribute, die in der erweiternden Klasse vorhanden sind ->ignore/überschreibe den Wert der erweiterten Klasse

Das folgende Beispiel soll verdeutlichen, was gemeint ist: Es wird eine vereinfachte Style Klasse definiert (Listing 3.4).

```
class Style(val name:String, val transparency:Double, val line-width:Int)
```

Listing 3.4: vereinfachte Style Klassendefinition

Nun wird eine Instanz *S1* der Modellklasse über die DSL definiert, welche sowohl einen Namen, als auch eine Angabe zur *transparency* hat (Listing 3.5).

```
style S1 {  
    transparency = 0.5  
}
```

Listing 3.5: Style S1

Eine weitere Instanz *S2* wird definiert, die ihrerseits über einen Namen und eine Angabe zur *line-width* verfügt. Außerdem soll *S2* die Instanz *S1* erweitern:

```
style S2 extends S1{  
    line-width = 10  
}
```

Listing 3.6: Style S2

Nach dem Parsen dieser Definitionen sollten Objekte folgender Struktur entstanden sein (siehe Listing 3.7).

```

{
  "name": "S1",
  "transparency": 0.5,
  "line-width": "notDefined"
}

{
  "name": "S2",
  "transparency": 0.5,
  "line-width": 10
}

```

Listing 3.7: *S1* und *S2* Objekte (aus Gründen der Lesbarkeit aufgeführt in der *Java Script Object Notation*)

Wie man sieht verfügt *S2* über eine *transparency* Angabe, die von *S1* übernommen wurde. *S1* hat von keiner Seite eine Angabe zur *line-width* bekommen. Unternimmt man an dieser Stelle Überlegungen zur Implementierung, wird man mit folgenden Fragen konfrontiert:

- Wie weiß ein Objekt, dass es das Attribut des Elternteils übernehmen oder überschreiben soll?
- Übernimmt *S2* tatsächlich alle Attribute des Elternteils oder hält es nur eine Referenz auf *S1*?

Die meisten der Attribute, die in der Definition einer Modellklasse berücksichtigt werden können, sind optional. Wie man bei der Definition von *S1* sieht, wurde keine Angabe zur *line-width* gemacht. Da die Definition jedoch auf den Konstruktor der Style Klasse abgebildet wird, muss irgendwo eine konkrete Angabe zu dem *line-width* Attribut erfolgen. Man könnte sich überlegen, jedem Feld der Modellklassen einen Defaultwert zuzuweisen. Das bedeutet, dass die *line-width* beispielsweise standardmäßig auf 10 gesetzt wird, wenn in der Definition keine Angabe darüber gemacht wird. Hat aber folglich jede Style Instanz zu jedem möglichen Attribut eine valide Angabe, fragt sich, woher später bekannt ist, ob auf Attribute einer Elterninstanz oder auf die eigenen zugegriffen wird. Immerhin gibt es zu jedem Feld der Instanz gültige Werte.

### 3.3.1 Kennzeichnung nicht definierter Felder in Modellklassen

Da Integer werte ausschließlich gültige Werte haben, ist es sinnlos 0 als *nicht gesetzt* zu verwenden und 1 - *MAX\_INT* als *gesetzte Werte*. Am Beispiel der *line-width* Angabe in *S1* wird also klar, dass bekannt sein muss, ob die Definition den Wert beschreibt oder nicht. Deshalb werden alle Felder der Modellklassen (Style, Shape, Connection, Diagram), die ein konkretes Attribut der Modellklasse darstellen, als sogenannte *Options* definiert. Options sind dem Paradigma der funktionalen Programmierung zu verdanken (z.B. in Haskell -> Maybe). Sie lösen ein Problem, welches in Java meist mit einer Exception gelöst wird. Eine Funktion gibt nicht immer für jedes Argument ein gültiges Ergebnis zurück (vgl.

Esser 2011, S.102). Im Falle der optional setzbaren Attribute der Modellklassen, bieten Options daher die passende Lösung. Man kann sie entweder auf *None* initialisieren oder auf ein *Some[T]*. Repräsentativ steht also in diesem Fall *None* für: „nicht gesetzter Wert“ (schau in den Elternklassen nach entsprechenden Werten) und *Some[T]* für: „gesetzter Wert“, ignoriere/überschreibe den Wert der Elternklasse.

### 3.3.2 Vererbungsmechanismus über Referenz auf Elterninstanzen

Als nächstes ist es derzeit auch noch unklar an welcher Stelle dieser Gedanke Anwendung findet. Es besteht einerseits die Möglichkeit, erst beim unmittelbaren Zugriff auf die einzelnen Felder zu prüfen, ob sie definiert wurden. Andererseits könnten jegliche Elterninstanzen schon bei der Erzeugung einer neuen Instanz nach fehlenden Werten durchsucht werden. Genauer ist also die Frage, ob Modellklassen eine zusätzliche Referenz auf Elternteile erhalten und nicht definierte Werte erst bei Bedarf in den Attributen der Elternliste gesucht werden oder ob Modellklassen schon bei der Instantiierung über sämtliche Attribute der Eltern Bescheid wissen, diese auf sich übertragen und anschließend nicht mehr auf ihre Eltern angewiesen sind. Beide Möglichkeiten bieten Vor- und Nachteile. Referenzen erlauben es, die Extraktionslogik der Attribute auf eine beliebige Stelle im Code auszulagern, je nach dem wann der Aufruf erfolgen soll.

### 3.3.3 Vererbungsmechanismus durch direkte Weitergabe der Attribute bei Instantiierung

Werden die Attribute der Elternklasse jedoch direkt bei der Objekterzeugung an die neue Instanz weitergegeben, hat dies zwei besondere Vorteile. Je nach Größe der Vererbungshierarchie ist es wesentlich performanter, die Werte im aktuell benutzten Element zu suchen, als über eine Referenz in den vielen Elternknoten ausfindig zu machen. Außerdem beseitigt diese Vorgehensweise die Rekursivität der Problemstellung.

Dazu ein stark vereinfachtes Beispiel: Werden mehrere Modellklassen definiert (siehe Listing 3.8), welche den jeweiligen Vorgänger erweitern und wird davon ausgegangen, dass die Attribute der zu erweiternden Instanzen direkt bei der Erzeugung übernommen werden,

```
style S1 {}
style S2 extends S1 {}
style S3 extends S2 {}
style S4 extends S3 {}
```

Listing 3.8: Definition einiger Styles, die vom jeweils nächsten Style erweitert werden

kann davon ausgegangen werden, dass *S2* jedes Element von *S1* erbt, *S3* jedes Element von *S2* erbt und somit auch implizit jedes Element von *S1* usw. Der

Reihe nach werden die einzelnen Styles erzeugt. Wird  $S_2$  erzeugt erbt es direkt alle Attribute von  $S_1$ , die es nicht selbst definiert (und somit gegebenenfalls überschreibt). Wird nun  $S_3$  erzeugt, erbt dieser Style nicht mehr direkt von  $S_1$ , da  $S_2$  garantiert alle Felder von  $S_1$  übernommen oder überschrieben hat. So muss bei der Vererbung also immer nur auf die nächst höhere Elterninstanz zugegriffen werden. Diese Eigenschaft lässt sich auf beliebig viele erweiternde Instanzen anwenden, wodurch in dem Beispiel  $S_4$  garantiert jede Eigenschaft von  $S_1$ ,  $S_2$  und  $S_3$  enthält, oder überschreibt. Da die Rekursivität der Problemstellung so entfällt und den Code (ausschlaggebender Faktor ist die Größe des durch Vererbung erstellten Baumes) performanter gestaltet, wird versucht die Vererbungshierarchie auf diese Weise aufzulösen.

### 3.4 Generatoren

Ausgehend davon, dass nun Strings eingelesen und in entsprechende Instanzen der Modellklassen umgewandelt werden können, ist der nächste Schritt über Generatoren den nun eigentlich gewünschten Code (Generat) zu erzeugen. Hierfür werden entsprechend der Modellklassen einige Generatorenklassen wie z.B. *StyleGenerator* erstellt, welche die Modellklassen in das gewünschte Format umwandeln. Die Generatorenklassen bilden die entstandenen Objektinstanzen auf *Javascript Code* ab.

# Kapitel 4

## Umsetzung

Basierend auf den Gedankengängen, die im Ansatz erläutert wurden, wird in diesem Kapitel die Umsetzung beschrieben. Hierbei werden zuvor getroffene Entscheidungen oftmals revidiert und neue Entschlüsse getroffen. Es wird mit Hilfe von Codebeispielen erläutert, was in dem entstandenen Programm passiert. Der komplette Quellcode zur Arbeit ist unter der Webadresse: <https://github.com/muellerjHTWGKonstanz/ScalaAsGeneratorLanguage.git> einzusehen.

### 4.1 Objektstruktur

#### 4.1.1 Modellklassen

Da bevor jeder anderer Logik zunächst die zu verwendenden Klassen definiert werden müssen, sollen Diese auch zuerst vorgestellt werden.

Die Modellklassen *Style*, *Shape*, *Connection* und *Diagram* sind entgegen des Ansatzes, diese als *case classes* umzusetzen, als normale Klassen implementiert worden. Dies hängt damit zusammen, dass die Parserlogik auf verschiedene Einzelparser aufgeteilt wurde. Entgegen der Objektstruktur, die durch die DSL in Xtext beschrieben wird, mussten einige Beziehungen geändert werden (siehe Abbildungen 4.1, 4.2, 4.3). Die Xtext Grammatik definiert zwar, welche Elemente nur Definitionsregeln sind und was in die Objektstruktur aufgenommen wird, jedoch liegt es letztlich beim Entwickler zu entscheiden, was in die endgültige Objektstruktur übernommen wird und was nicht.

#### 4.1.2 Besonderheiten bei Style

Bei der Umsetzung der einzelnen Modellklassen aus der Xtext Grammatik in die Objektstruktur in Scala, kam es regelmäßig zu Situationen, in denen die Objektstruktur nicht 1:1 auf Scala übernommen werden konnte, da die verschiedenen Sprachen verschiedene Features bedienen, oder eben nicht (z.B. Xtext Schlüsselwort dispatch). Außerdem galt es bestehende komplexe Strukturen möglicherweise vereinfacht umzusetzen. Beispielsweise enthält die Definition der *Style*

klasse in Xtext ein Style Layout, welches die eigentlichen Attribute der Style klasse auflistet (siehe Listing 4.1).

```
Style:
    {Style}
    [...]
    layout=StyleLayout
    [...]
;

StyleLayout:
    {StyleLayout}
    (
        ("transparency" "=" transparency=DOUBLE)? &
        [...]
    )
;
```

Listing 4.1: stark vereinfachter Auszug aus der Xtext Grammatik, die Styles beschreibt

In diesem Fall wurden alle Attribute eines Styles direkt in der Style klasse aufgelistet. Dies vereinfacht einerseits die Objektstruktur, außerdem erleichtert es den Zugriff auf die Attribute, da die Navigation durch die referenzierte Instanz eines Layouts entfällt. Des weiteren definiert die Xtext Grammatik nicht eindeutig um welche Abstraktionsform es sich bei manchen Definitionen handelt (siehe Listing 4.2).

```
ColorOrGradient: Color | Transparent | GradientRef
```

Listing 4.2: Auszug aus der Style.xtext Grammatik, Definition der ColorOrGradient

Also kann an Stelle eines *ColorOrGradient* sowohl eine *Color*-, *Transparent*- oder *GradientRef* Definition stehen. Für das Einbetten in eine Objektstruktur sind diese Regeln nicht ideal, da *Color*, *Transparent* und *GradientRef* keine gemeinsamen Eigenschaften haben. Sie lassen sich zwar unter einer gemeinsamen Abstraktion zusammenfassen, dies bereitet jedoch später im Generator Probleme. *Highlighting* Attribute sind beispielsweise vom Typ *ColorOrGradient*. Später im Generator werden für verschiedene Untertypen von *ColorOrGradient* einheitliche *Getter* Methoden definiert (dies funktioniert in Xtext über eine sogenannte *dispatch* Kennzeichnung), jedoch geht *Gradient* dabei leer aus. Wird also ein *Gradient* definiert, muss der Entwickler blind vertrauen, dass es nicht an der falschen Stelle eingesetzt wurde. Gibt man einem *highlighting* Attribut ein *Gradient*, was laut Abstraktion möglich ist, wird in diesem Fall die Methode *createColorValue* in einem *Gradient* gesucht, welche diese Methode **nicht** beschreibt. Entwickler müssten also jedes mal wenn die entsprechende Methode aufgerufen wird, zunächst prüfen um was für eine konkrete Unterklasse von *Co-*



*lorOrGradient* es sich handelt. Für die Vereinheitlichung der Klassen wurden Traits benutzt. Ausdrücke wie in Listing 4.3

```
ColorOrGradient: Color | Transparent | GradientRef;
ColorWithTransparency: Color | Transparent;
```

Listing 4.3: Auszug aus der Style.xtext Grammatik, Color kann als ColorOrGradient oder als ColorWithTransparency benutzt werden

zeigen, dass Xtext eine Art der Mehrfachvererbung zulässt. *Color* kann sowohl als *ColorOrGradient*, als auch *ColorWithTransparency* benutzt werden. Traits haben die Möglichkeit in eine Klasse “hineingemischt” (vgl. Esser 2011, S.191) zu werden und sollen in diesem Sinne in der Regel weitere Eigenschaften hinzufügen. Hierüber wird das Problem der vermeintlichen Mehrfachvererbung gelöst, da beliebig viele Traits in die Klasse hineingemischt werden können. Einige Elemente wie zum Beispiel Farbkonstanten (“light-orange“ -> LIGHT\_ORANGE, “blue“ -> BLUE), wurden nicht wie in Xtext als Enumeration verwirklicht, sondern als *case objects* (siehe Listing 4.4).

```
case object WHITE extends ColorConstant {def
  getRGBValue = ""#ffffff"" }
case object LIGHT_GRAY extends ColorConstant {def
  getRGBValue = ""#d3d3d3"" }
case object GRAY extends ColorConstant {def
  getRGBValue = ""#808080"" }
case object DARK_GRAY extends ColorConstant {def
  getRGBValue = ""#a9a9a9"" }
...
```

Listing 4.4: Auszug aus Code ColorConstants als case objects

Da Farbkonstanten (*ColorConstant*) so von Superklassen erben können, ist es möglich einheitliche Methoden für sie zu definieren. Beispielsweise ist die Methode *createOpacityValue* in einer Superklasse einheitlich beschrieben worden (siehe Listing 4.5).

```
trait ColorOrGradient{
  def getRGBValue:String
  def createOpacityValue:String = "1.0"
}
```

Listing 4.5: Auszug aus Code Trait ColorOrGradient

Da ColorOrGradient ein Trait ist, erlaubt es (anders als Javas *Interfaces*), die Methoden vorzuimplementieren. So ist automatisch jede Farbkonstante bedient und es muss nicht erneut wie in Xtext auf eine Weise wie *dispatch* gearbeitet werden.

In Abbildung 4.1 ist eine vereinfachte Version der entstandenen Objektstruktur zu sehen, die die wichtigsten Abhängigkeiten der Style Modellklasse darstellt.

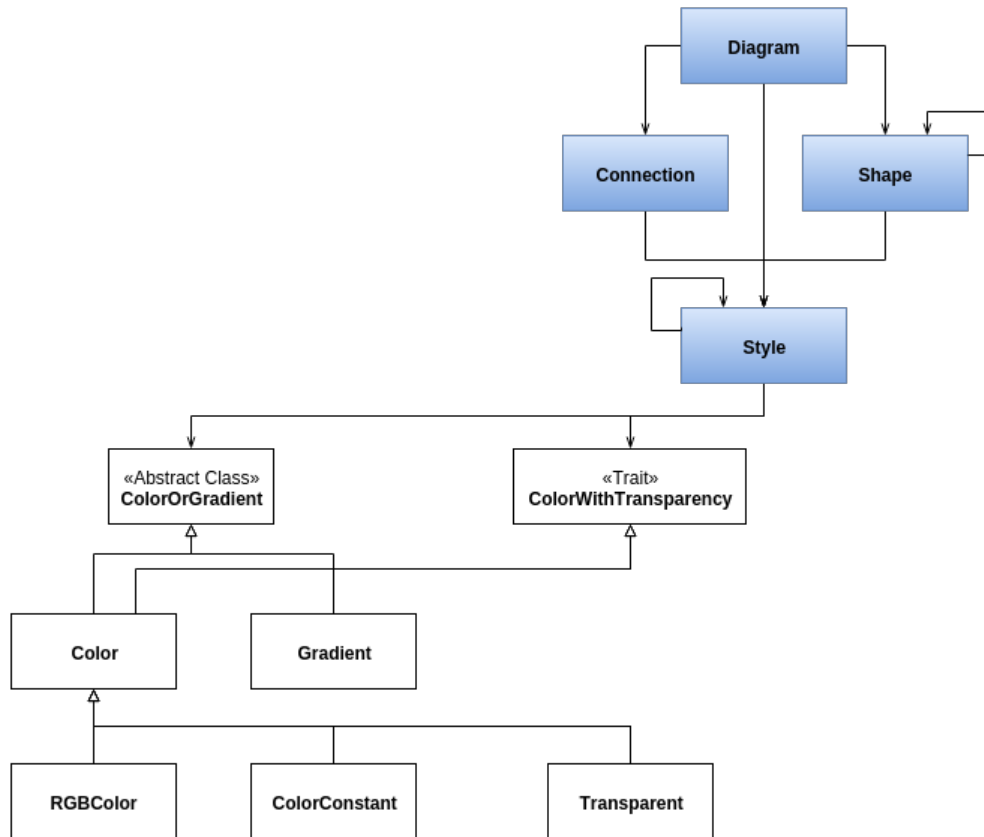


Abbildung 4.1: Vereinfachte Objektstruktur. Im Fokus Style und seine wichtigsten Abhängigkeiten.

### 4.1.3 Besonderheiten bei Shape

Für Shapes muss zunächst ein Namensproblem aufgelöst werden. Mit einer Shape wird der Wrapper definiert, der sowohl Metainformationen zu der Shape enthält als auch die **eigentlichen Shapes**, in diesem Fall geometrische Figuren (*GeometricModel*), enthält. Ist also die Rede von GeometricModels, sind die tatsächlichen Implementierungen beispielsweise einer Ellipse gemeint. Ist die Rede von einer Shape, ist die Wrapper Klasse gemeint, welche beispielsweise Ellipsen beinhalten kann. Da die verschiedenen geometrischen Figuren einheitlich behandelt und verwaltet werden können müssen, bildet GeometricModel ihre gemeinsame Abstraktion (siehe Abbildung 4.2). Die konkreten geometrischen Figuren haben zwar untereinander hin und wieder Gemeinsamkeiten, doch es lassen sich tatsächlich keine allgemeinen Ähnlichkeiten feststellen. Beispielsweise haben einige zwar Angaben zur Größe, jedoch nicht alle. Aus diesem Grund definiert die abstrakte Klasse GeometricModel lediglich eine Gemeinsamkeit, die für die Distribution von vererbten Informationen wichtig ist: Die Referenz

auf ein eventuell existierendes Elternteil. Ein Rechteck, kann ein Polygon beinhalten, welches wiederum ein weiteres Rechteck beinhalten kann (siehe Listing 4.6).

```
ellipse style StandardStyle {  
    rectangle {  
        ...  
    }  
}
```

Listing 4.6: Beispiel Ellipse die einer Styleregeln unterliegt und ein Rechteck beinhaltet

In diesem Falle muss garantiert werden können, dass auch das verschachtelte Rechteck über die Style-Attribute der Ellipse verfügt. Die Layouts der einzelnen geometrischen Formen wurden als Traits realisiert. Der große Vorteil der Traits als *rich-Interfaces* liegt darin, dass mehrere Eigenschaften (Traits) geerbt werden können. So kann jede geometrische Figur von der abstrakten Klasse `GeometricModel` erben und gleichzeitig die Eigenschaft eines spezifischen Layouts besitzen. Die Layouts selber haben wiederum nur eine Gemeinsamkeit, nämlich die Möglichkeit einen Style zu referenzieren. Diese Eigenschaft wurde in dem Trait `Layout` bestimmt, von dem **jedes** weitere Layout erbt (siehe Abbildung 4.2). Da die geometrischen Figuren nun keine Referenz mehr auf Layout Objekte haben, sondern entsprechende Felder selber erben, ist auch hierbei der Navigationsaufwand deutlich verringert. Wenn ein `roundedRectangle` zuvor auf seine Positionskoordinaten zugreifen wollte, gelang dies wie in Listing 4.7.

```
val x_coord = rr.layout.common.x
```

Listing 4.7: Beispielnavigation durch komplexe Objektstruktur

Durch die Vereinfachung der Objektstruktur und die Umsetzung der Layouts als Traits, können diese Informationen direkt abgerufen werden, sind aber logisch weiterhin von der erbenden Klasse getrennt und können einheitlich gewartet werden. In Abbildung 4.2 ist eine vereinfachte Version der entstandenen Objektstruktur zu sehen, die die wichtigsten Abhängigkeiten der Shape und Connection Modellklassen darstellt.

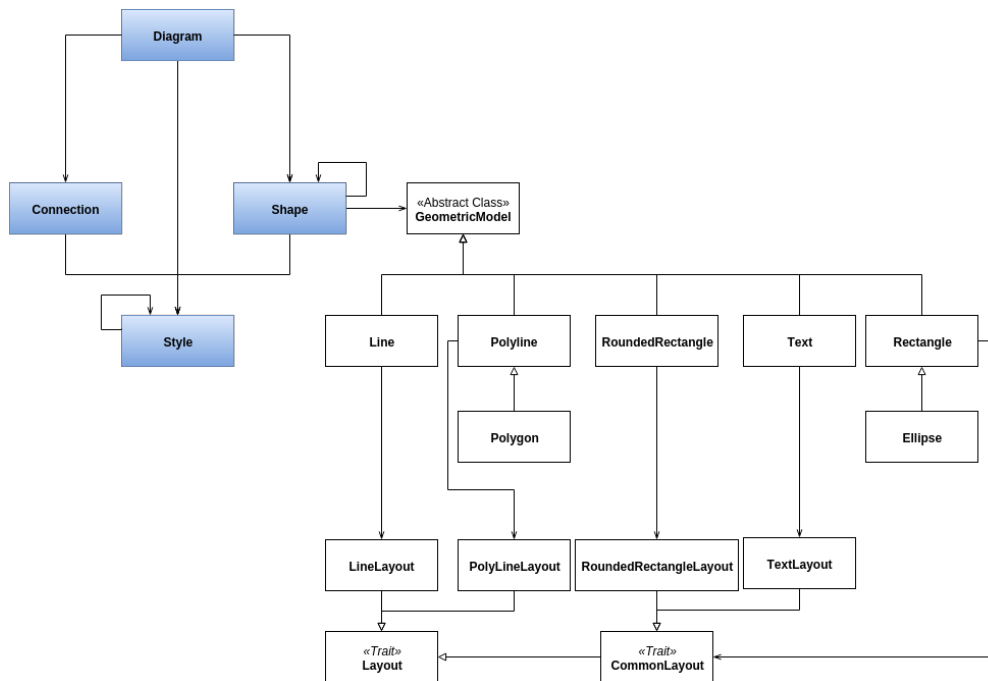


Abbildung 4.2: Vereinfachte Objektstruktur. Im Fokus Shape und seine wichtigsten Abhängigkeiten.

#### 4.1.4 Besonderheiten bei Diagram

Die DSL, in der Diagrams geparkt werden, verknüpft das Metamodell und die Modellklassen. Die Definition eines Diagrams sieht (vereinfacht) aus wie in Listing 4.8.

```

diagram diagramName for metamodelElement (style :
  styleName) {
    ...
}

```

Listing 4.8: Beispielhafte Diagram Definition (vereinfacht)

Diagrams und einige seiner Attribute verweisen auf ein Metamodell oder dessen Felder, aus dem MoDiGen Projekt. Da das Metamodell zur Zeit der Entwicklung noch nicht bereit war, wurden diese Elemente mit Mockups versehen (siehe Listing 4.9). Ein Auszug aus dem SprayParser:

```

trait MclassMock {
  var name:String
  def superTypeOf(childClass:MclassMock):Boolean = true
  ...
}

```

Listing 4.9: Auszug aus Code MclassMock über dass eigentliche MClass Elemente gemockt werden

Bemerkenswert ist bezüglich des Zugriffs auf das Metamodell an dieser Stelle, dass Zugriffe auf einzelne Elemente des Metamodells im Gegensatz zu Xtext sehr einfach gestaltet werden können. Dieser Vorteil gegenüber Xtext rührt nicht allein durch Scala, sondern der Entscheidung Metamodelle in der *JavaScript Object Notation* in einer *NoSQL* Datenbank zu hinterlegen. Xtext benutzt *XML (EXtensible Markup Language)* Dateien als Speichermedium. Zugriff auf spezifische Elemente erfolgt in Xtext über das Laden der gesamten XML Datei (vgl. Gunderloy and Tittel 2004, S.58-60). Da die Entitäten der Metamodelle in MoDiGen einzeln im *JSON* Format in eine *NoSQL* Datenbank eingespeist werden, ist es auch möglich, einzeln auf entsprechende Elemente zuzugreifen. Diese Möglichkeit des Datenzugriffs vermeidet (abhängig von der Dateigröße) unter Umständen riesigen Overhead und ist auf Grund der geringeren Bandbreitenauslastung auch wesentlich performanter. In Abbildung 4.3 ist eine vereinfachte Version der entstandenen Objektstruktur zu sehen, die die wichtigsten Abhängigkeiten der Diagram Modellklasse darstellt.

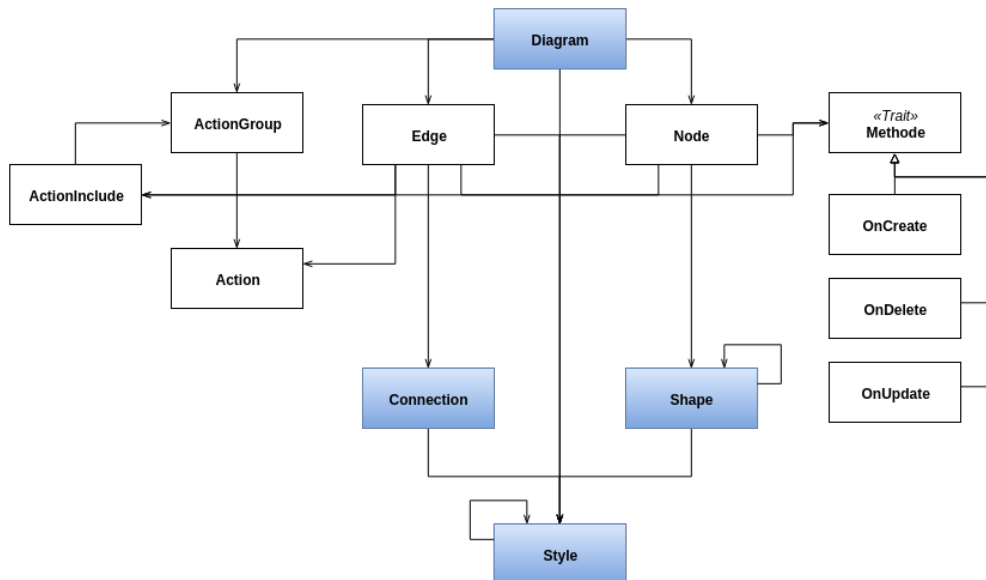


Abbildung 4.3: Vereinfachte Objektstruktur. Im Fokus Diagram und seine wichtigsten Abhängigkeiten.

#### 4.1.5 Vererbung/ClassHierarchy

Nachdem die Modellklassen definiert wurden, müssen diese wie bereits beschrieben auch hinsichtlich der Vererbungshierarchie verwaltet werden können. Entsprechend des Ansatzes wurde - wie ursprünglich geplant - eine generische Klasse namens *ClassHierarchy* erstellt, welche sowohl Knoten definiert, als auch eine Map führt, die Namen (Strings) auf entsprechende Knoten abbildet. Da diese Klasse im Stande sein soll, die Namen, also die tatsächlichen String Attribute der Style- und Shape Klassen abzubilden, wurde in der Typparameterliste ein *Viewbound* definiert (siehe Listing 4.10).

```
sealed class ClassHierarchy[T <% {val name:String}](  
    rootClass:T){
```

Listing 4.10: Auszug aus Code Viewbound. Implizite Umwandlung auf `val name:String`

Die Kennzeichnung `X <% Y` sagt aus, dass der übergebene Typ `X` implizit auf den Typ `Y` umgewandelt werden können muss (vgl. Braun 2011, S.158). In Diesem Fall wurde ein anonymer Typ erstellt, der ein Feld `name:String` besitzt. Wenn der angegebene Typ also über ein solches Feld verfügt, kann er implizit darauf abgebildet werden und wird akzeptiert. Des weiteren wurden mehrere *Apply*-Methoden definiert, die den Umgang mit *ClassHierarchy* erleichtern sollen (siehe Listing 4.11).

```
def apply(parent:Node, className:T) =  
    parent inheritedBy className  
def apply(parent:T, className:T) =  
    nodeView(parent.name) inheritedBy className  
def apply(parent:String, className:T) =  
    nodeView(parent) inheritedBy className  
def apply(className:T) =  
    nodeView(className.name)  
def apply(className:String) =  
    nodeView(className)
```

Listing 4.11: Auszug aus Code Apply Methoden, die die Benutzung von *ClassHierarchy* erleichtern

Apply-Methoden sind spezielle Methoden und können ohne den eigentlichen Methodennamen aufgerufen werden. Die Argumente werden einfach nach der Instanz in Klammern angegeben (vgl. Esser 2011, S.74). Dies ist sehr komfortabel. Eltern- und Kindbeziehungen werden in den Knotenklassen gespeichert. Leider ist *ClassHierarchy* so jedoch, wie sich herausstellt, nicht wie gewünscht einsetzbar. Wird ein neuer Style oder eine neue Shape erzeugt, müssen die Eigenschaften der Elternknoten verfügbar sein, noch bevor der/die eigentliche Style/Shape instantiiert wurde, da diese Eltern und deren Felder noch vor dem Instantiieren des neuen Objekts aufgelistet und abgerufen werden. Das *Companion Objekt* der entsprechenden Klassen sammelt erst alle Informationen an

neu geparsten und an geerbten Attributen und gibt diese dann an den richtigen Konstruktor weiter (siehe Abschnitt 4.3). Bevor also ein Style/Shape instantiiert ist, kann in *ClassHierarchy* auch nicht danach gesucht werden, es existiert ja noch nicht. Entsprechende Elternknoten des noch nicht existierenden Knotens sind so ebenfalls nicht ausfindig zu machen. Dieser Denkfehler reduziert die Brauchbarkeit der *ClassHierarchy*-Klasse auf die Map[String, T], welche diese noch enthält. Da die Namen der Elternklassen über die Definition der Modellklasse mitgeliefert werden, werden die Entsprechenden Elternteile über die Abbildung String=>T der Map gefunden. Die komplette, gebrauchte Funktionalität der *ClassHierarchy* ist also bereits durch eine herkömmliche Map der scala.collections beschrieben. Lediglich die zusätzlichen apply-Methoden der *ClassHierarchy*, welche den Gebrauch der Map vereinfachen, sorgen derzeit dafür, dass die Klasse einen Nutzen erfüllt. Schließlich behalten die erzeugten Styles/Shapes eine Liste mit ihren Eltern. Da sie in sich alle Merkmale der Superklassen behalten, wäre dies nicht einmal nötig. Jedoch bleibt es zu Testzwecken nützlich eine solche Referenz zu haben, um schnell überprüfen zu können welche Werte übernommen wurden, welche nicht und warum nicht (z.B. *latest-bound principle*).

## 4.2 Implementierung der Parserklassen

Sind die Beziehungen unter den Modellklassen definiert, muss als nächstes dafür gesorgt werden, dass entsprechende Klassen auch instantiiert werden können. Da die benötigte Information für die Erstellung einer Modellklasse in Stringform bereitgestellt wird, ist also nun ein Parser einzurichten, der die benötigte Umwandlung der in Stringform enthaltenen Attribute zu den richtigen Datentypen vollziehen kann.

### 4.2.1 Factory Klassen

Der Parser muss, dem Metamodell entsprechend, unter anderem sowohl Style-, Shape-, als auch Diagram Definitionen verarbeiten können (siehe Abbildung 1.2). Entsprechend wurden Diagram.scala, Shape.scala und Style.scala erstellt. Zunächst waren diese Klassen als *case classes* geplant, um bestimmte Methoden vorimplementiert benutzen zu können. Allerdings wurde schlussendlich entschieden, Teile der Parserlogik auszulagern. Gemäß dem *Single Responsibility Principle* (vgl. Martin 2008, S.181), lag es nahe sowohl einen Style-, Shape- und Diagramparser zu haben, auf die der eigentliche Parser zurückgreifen kann. Im Idealfall würden diese spezialisierten Parser wie *Factorys* agieren, um die zugehörige Klasse möglichst entkoppelt instantiiieren zu können. Die *Factorys* wurden als *Companion Objekte* realisiert, da diese sich hierfür besonders eignen. Diese Absicht kreuzt sich leider mit den *case classes*, da diese bereits implizit über ein *Companion Objekt* verfügen. Diese *Companion Objekte* können nicht partiell überschrieben werden. Jede der (normalen) Modellklassen hat nun ein *Companion Objekt*, welches eine *parse* Methode enthält, um seine *Companion Klasse* zu instantiiieren. Die *Companion Objekte*, können hierbei als *Factory*

angesehen werden, da ihre *apply*-Methoden den Parse Vorgang ebenfalls initiieren und anschließend eine neue Instanz ihrer *Companion Klasse* zurückliefern (vgl. Esser 2011, S.80). Eine Modellklasse kann also erzeugt werden, indem ohne *new* das Companion Objekt mit der entsprechenden Argumentenliste aufgerufen wird. Die Fabrikmethode bekommt Argumente als Tupel aus Strings Form (z.B. (“line-width“, “4“) -> Integer 4 wird aufgelöst und anhand des Attributnamens an der richtigen Stelle im Konstruktor platziert) und der eigentliche Konstruktor nur aufgelöste Werte als Argumente akzeptiert. Das oben beschriebene Prinzip, der Aufteilung der Parserlogik in mehrere kleine Parser, wird konsequent durchgeführt. So haben eigene Datentypen wie z.B. *GradientAlignment*, oder *Anchor* ebenfalls kleine Parsereinheiten in ihren Companion Objekten. Um sicherzustellen, dass auf eigene Faust keine Modellklasse erzeugt werden kann, wird der Konstruktor auf *private* gesetzt und die Klasse mit *sealed* gekennzeichnet. So ist es ausschließlich über das Companion Objekt möglich eine Modellklasse zu erzeugen. Dem Companion Objekt ist der Zugriff auf den privaten Konstruktor möglich, da es Zugriff auf alle privaten Member der *Companion Klasse* hat. Ansonsten könnten im schlimmsten Fall sogar abstrakte Klassen anonym instantiiert werden (siehe Listing 4.12).

```
val gradient_alignment = new GradientAlignment() {}
```

Listing 4.12: Beispiel zur Instantiierung der abstrakten Klasse GradientAlignment

Das Schlüsselwort *sealed* verhindert, dass die gekennzeichnete Klasse außerhalb ihrer Quelldatei erweitert werden kann. So wäre es erneut möglich die Klasse um eine Ecke selber zu instantiieren (siehe Listing 4.13).

```
class AnotherGradient extends GradientAlignment
&
val anonymous_gradient = new AnotherGradient()
```

Listing 4.13: Beispiel zur Instantiierung der abstrakten Klasse GradientAlignment über Vererbung

Da davon ausgegangen werden kann, dass die Modellklassen nur über geparsen Input erstellt werden können, ist es sicherer auch ihre Instantiierung **ausschließlich** den Parsern zu überlassen. Deshalb sollten alle Modellklassen auf diese Weise definiert werden.

## 4.2.2 Problematik der rekursiven regulären Ausdrücke

Das tatsächliche Parsen der Strings, erfolgte zunächst über reguläre Ausdrücke. Warum dies in der Umsetzung scheitert und wie das Problem zu lösen ist, wird in diesem Abschnitt erläutert.



Zunächst erscheint es einleuchtend die Strings über *Regular Expressions* einzulesen und auszuwerten. Ist die zugrundeliegende Grammatik simpel genug, funktioniert dies auch einwandfrei. Simple String *match*-Funktionen lassen sich in Scala sehr viel komfortabler anwenden als beispielsweise in Java. Nun wurde zunächst der Style Parser komplett und fehlerfrei fertig gestellt. Die bestehende Lösung definierte viele undurchsichtige Regeln, nicht zuletzt da komplexe reguläre Ausdrücke einfach nicht schön zu lesen sind. Entsprechend ist die Skalierbarkeit einer solchen Lösung nicht sehr effizient, da um eine ohnehin schon komplexe Regel zu erweitern, erst einmal die bestehende Regel verstanden werden muss, was wie bereits erwähnt nicht trivial ist. Dies kann einige Zeit in Anspruch nehmen. Beim erstellen der Regeln für die Shape Klassen stießen die regulären Ausdrücke endgültig an ihre Grenzen. Vor Allem die Anforderung der DSL, ineinander geschachtelte geometrische Figuren abbilden zu können, ist mit regulären Ausdrücken nicht machbar. Eine Shape Definition wie in Listing 4.14.

```
shape exampleShape {
  ellipse {
    size (width=50, height=50)
    ellipse {
      size (width=38, height=38)
    }
  }
}
```

Listing 4.14: Beispielhafte Shapedefinition

Die Shape in Listing 4.14 ist einerseits sehr komplex, da man um sie parsen zu können sehr komplizierten und unschönen Code erzeugen muss. Das Problem ist, dass beim Parsen zwar ein „size(width=50, height=50)“ Attribute ermittelt werden kann, die Zuweisung zur richtigen Ellipse aber höchstens auf einer Indexvariable basieren könnte, die die *size* an Stelle *n*, der geometrischen Figur an Stelle *n* zuweist. Doch hierbei müssten zahlreiche Fälle beachtet werden, zumal nicht für jede geometrische Figur jede mögliche Angabe in der Definition auch erfolgen muss. Das weitaus größere Problem ist, dass normale reguläre Ausdrücke an rekursiven Ausdrücken scheitern, da Scala (Stand Version 2.11.7) keine rekursiven regulären Ausdrücke unterstützt. Setzt die Grammatik also voraus, dass ein Element **beliebig** oft in sich selbst gekapselt werden kann (z.B. Listing 4.15),

```
Babuschka {
  Babuschka {
    Babuschka {}
  }
}
```

Listing 4.15: Beispiel für eine rekursive Grammatik anhand der russischen *Babuschka* Puppen

sind reguläre Ausdrücke dem Problem nicht mehr gewachsen. In diesem Fall muss auf eine komplexere Technologie zurückgegriffen werden.

### 4.2.3 Lösung Parser Combinators

Hierbei bieten sich die *Parser Combinator* Klassen an. Wie der Name schon vermuten lässt, kann man hier Parser zusammenschalten, um so komplexe Ausdrücke auswerten zu können. Hierfür werden Methoden definiert, welche als Rückgabewert einen *Parser[T]* liefern. Für *Parser* sind unter anderem die Methoden  $\sim$ ,  $\sim>$  und  $<\sim$  angegeben, welche sich wie normale Operatoren anfühlen und verwenden lassen. Braun hat dies in seinem Werk *SCALA Objektfunktionale Programmierung* wie folgt einleuchtend erklärt (Braun 2011, S.182):

- $a \sim b$  erzeugt einen Parser, der zuerst  $a$ , dann mit  $b$  parst und das Gesamtergebnis zurückgibt.
- $a \sim> b$  parst genauso wie  $a \sim b$ , gibt aber nur das Ergebnis des Parsers  $a$  zurück.
- $a <\sim b$  gibt analog nur das Ergebnis von  $b$  zurück.

Zunächst ein einfaches Beispiel in der interaktiven scala *REPL* (siehe Listing 4.16), in dem ein Parser namens *NameParser* erstellt wird. Für diesen wird eine Methode *name* definiert, welche aus einem String „hello my name is <name>“ den entsprechenden Name filtert und zurück liefert.

```
scala> import scala.util.parsing.combinator.
      JavaTokenParsers
import scala.util.parsing.combinator.JavaTokenParsers

scala> class NameParser extends JavaTokenParsers {
      | def name = "hello my name is" ~> ident ^^ {_.
      |   toString}
      | }
defined class NameParser

scala> val p = new NameParser()
p: NameParser = NameParser@2ecaa79e

scala> p.parse(p.name, "hello my name is Julian")
res0: p.ParseResult[String] = [1.24] parsed: Julian
```

Listing 4.16: Einfaches Beispiel für einen Namens Parser

Mit relativ wenig Aufwand können sogar sehr komplexe Grammatiken beschrieben werden (siehe Listing 4.17).

```
scala> import scala.util.parsing.combinator.
      JavaTokenParsers
```

```
import scala.util.parsing.combinator.JavaTokenParsers

scala> class BabuschkaParser extends JavaTokenParsers{
      def babuschka:Parser[String]=("babuschka" ~>
        ident) ~ ("{" ~> (babuschka|")") <~ "}" ^ ^ {_.
        toString}
    }
defined class BabuschkaParser

scala> val p = new BabuschkaParser()
p: BabuschkaParser = BabuschkaParser@23a1ef14

scala> p.parse(p.babuschka, "babuschka b1{ babuschka b2 {
      babuschka b3 {babuschka b4{}}}}")
res3: p.ParseResult[String] = [1.36] parsed: (b1~(b2~(b3~
      (b4~)))
```

Listing 4.17: Einfaches Beispiel zum Parsen einer rekursiven Grammatik

Umgesetzt wurden so beispielsweise die rekursiven *GeometricModels*- geometrische Figuren - welche in sich gekapselt erzeugt werden können (siehe Listing 4.18).

```
private def geoModel: Parser[GeoModel] =
  geoIdentifier ~
  (((("style" ~> ident)?) <~ "{" ) ~
  rep(geoAttribute) ~
  (rep(geoModel) <~ "}") ^ ^ {
    case name ~ style ~ attr ~ children =>
      GeoModel(name, style, attr, children, cache)
  })
```

Listing 4.18: Rekursive Methode zum Parsen geometrischer Figuren

Wie man sieht ist die Methode *geoModel* rekursiv, da sie sich selbst aufruft. Doch zunächst erst ein paar einfachere Parser.

#### 4.2.4 Prinzip der Parserlogik

Die gewählte Technologie zum Parsen sind also *Parser Combinators*, über welche komplexe Problemstellungen gelöst werden können. Wie die einzelnen Parser untereinander assoziiert sind, was ihre jeweiligen Aufgaben sind und welche Gedanken dahinter stecken wird nun beschrieben.

Die grundlegende Parserinstanz, mit der auch der Benutzer schlussendlich interagiert ist der *SprayParser*. Wie bereits erwähnt, sollte dieser Parser möglichst generisch funktionieren, um das Parsen der verschiedenen Modellklassen im Idealfall möglichst einheitlich gestalten zu können (siehe 3.2.1). Konkret heißt das, dass versucht wurde so wenig expliziten Inhalt der Modellklassen abzufragen

wie möglich. Anstatt eine Regel aufzustellen, die also beispielsweise direkt nach einem *line-width* Attribut sucht, wurden Regeln gewählt, die allgemeiner auf die Attribute eingehen (siehe Listing 4.19).

```
def attributePair: Parser[(String, String)] =
    variable ~ arguments ^^ { case v ~ a => (v, a) }
def variable: Parser[String] =
    "[a-züäöA-ZÜÄÖ]+([-_][a-züäöA-ZÜÄÖ]+)*".r <~ "\\s*".r
def arguments: Parser[String] =
    argument_classic | argument_advanced_explicit |
    argument_advanced_implicit | argument_wrapped
```

Listing 4.19: Auszug aus Code beispielhafte Regeln, die verallgemeinerte Attribut- Wert Tupel parsen

Es wird also zunächst definiert, dass beim Parsen von Attributen ein *attributePair* erwartet wird. Dieses *attributePair* besteht aus einer *variable* und (*~*) *arguments*. Eine Variable wird in der nächsten Regel bzw. Methode definiert. Die vielen verschiedenen Möglichkeiten den Variablenwert zu ermitteln führt darauf zurück, dass dieser Teil nun eben sehr variabel ist. An dieser Stelle kann jeder dem System bekannte Datentyp folgen. Während ein *argument\_classic* erwartet, dass ein '=' Zeichen und anschließend irgendein Int, double oder String folgt, deckt ein *argument\_advanced\_explicit* die Möglichkeit eines Arguments wie in Listing 4.20.

```
someSize ( width = 10, height = 12)
```

Listing 4.20: Beispiel für eine Attributsdefinition mit expliziter Parameterbenennung

Da diese und weitere Parserregeln öfter gebraucht werden, sind diese in einem Trait namens *CommonParserMethods* zusammengefasst. Die *Companion Objekte* der Modellklassen, die gleichzeitig den Parser bilden, erweitern diesen Trait. Blickt man zurück auf *attributePair*, sieht man außerdem, dass diese Methode ein *Tupel* zurück gibt, welches eben zum einen den Variablennamen und außerdem den Wert zurückliefert. Da an dieser Stelle des Parsers noch keiner der Werte in seinen eigentlichen Typ konvertiert wird, kann nun auch ein einheitlich typisiertes Set an *Tupel[String, String]* an die tiefer liegenden Parser weitergegeben werden. Wird also gerade beispielsweise ein Style geparsed, werden die erhaltenen *Tupel* nun an das Companion Objekt *Style* übergeben, das sich selber um seine Attribute zu kümmern hat. Bisher wurde also nur darauf geachtet, dass der Inputstring möglichst mit keiner Konvention der zugrundeliegenden Grammatik bricht, außerdem wird so ein einheitliches Medium garantiert, nach dem sich die spezialisierten Parser der Modellklassen richten können. Da diese grundlegenden Regeln noch keine konkreten Instanzen erzeugen, sondern nur dem Identifizieren der einzelnen Wörter der DSL und der Bereitstellung einheitlicher Datenformate dient, bildet diese logische Einheit (SprayParser) den *Lexer* der Anwendung (vgl. Aho 2008, S.135). Da sich die spezialisierten Parser nun ihr Inputformat teilen, können sie folglich auch mit denselben Methoden weiterarbeiten. So muss das Rad nicht für jeden Parser neu erfunden werden.

Ab jetzt passiert im spezialisierten Style Parser die eigentliche Magie. Bevor die übergebenen Werte in den Tupeln nun ihre endgültige Form erhalten, widmen sich die *Factorys* zunächst der Vererbung. Dazu mehr in Abschnitt 4.3. Nachdem die vererbten Werte berücksichtigt sind, werden nun die Tupel mit den Attributen auseinandergenommen und vordefinierten Variablen zugewiesen, welche später an den Konstruktor der Modellklasse weitergegeben werden. Die Factory erhält also ein Set mit Attributen. Diese Attribute sind in Form eines `Tupels[String, String]` und enthalten jeweils den Namen des zuzuweisenden Feldes und den Attributswert (aber immer in String Form). Um Die Werte aufzulösen und zuzuweisen, wird eine mächtige Technik namens *Pattern Matching* eingesetzt (siehe Listing 4.21).

```
attributes.foreach{
  case ("description", x) =>
    description = Some(x)
  case ("transparency", x) =>
    transparency = ifValid(x.toDouble)
  case ("line-style", x) =>
    line_style= LineStyle.getIfValid(x)
  ...
}
```

Listing 4.21: Auszug aus dem Code zum Pattern Matching

So wird immer zunächst geprüft um welches Feld es sich handelt. Anschließend wird der String falls nötig auf den gewünschten Typ konvertiert und zugewiesen. Da Fehleingaben wie Buchstaben an Stelle einer erwarteten Zahl möglich sind, wird über eine Hilfsmethode *ifValid* (siehe Listing 4.22) sichergestellt, dass es nicht zum Programmabsturz kommt, sondern im Fehlerfall ein Defaultwert eingesetzt wird.

```
def ifValid[T](f: => T):Option[T] = {
  var ret:Option[T] = None
  try { ret = Some(f)
    ret
  }finally {
    ret
  }
}
```

Listing 4.22: Auszug aus dem Code Methode ifValid

Die Methode *ifValid* erwartet einen Typparamter und einen Codeblock, der ein Ergebnis entsprechend des Typparameters zurückliefert. Der gewünschte Codeblock wird in einer sicheren *try* Umgebung ausgeführt. Bei erfolgreicher Ausführung wird das Ergebnis des Codeblocks zurückgeliefert, ansonsten *None*. Der Aufruf von *ifValid* kann wie man in Listing 4.21 sieht, aber sogar ohne Typparameter erfolgen, da Scala diesen auch hier anhand der Zuweisung erkennen

kann und ihn inferiert (Scala ist klasse!). Im Falle der Styles werden größtenteils nur primitive Datentypen geparsed, daher ist das Konvertieren der Werte in die eigentlichen Datentypen auch eher einfach. Im Falle der Shape Modellklasse werden unter anderem mehrere ineinander gekapselte geometrische Formen geparsed. Beim Parsen der Ellipsen, Rechtecke etc. sind mehrere Parser beteiligt. Beim Pattern Matching des Shape Companion Objekts wird daher in der Regel auf weitere spezialisiertere Parser verwiesen (siehe Abbildung 3.2).

## 4.3 Vererbung

Da Styles und Shapes auch mit einem *extends* Schlüsselwort definiert werden können, muss auch ein Mechanismus vorhanden sein, der sich um die Vererbungshierarchie kümmert und die Felder der erweiterten Klassen abrufen kann. Bisher wurde die Vererbung über eine Referenz zum Elternknoten geregelt. Wenn also ein gefragter Wert in der aktuellen Klasse nicht gefunden wurde, suchte man rekursiv in den Elternklassen danach. Dies erfordert nun einerseits bei größeren Hierarchien einige Zeit, außerdem eventuell komplizierte rekursive Funktionen. Ein leichter und performanter Weg wurde darin gefunden, die Werte eines Elternknotens direkt beim Erzeugen der neuen Instanz auf das Kind zu übertragen. In dem Companion Objekt von Style/Shape, in dem die Attribute für einen neuen Style aufgelöst werden, wird hierfür zuerst die Liste an Eltern erstellt (siehe Listing 4.23).

```
val extendedStyle:List[Style] = parents.getOrElse(List())
    .foldLeft(List[Style]())((styles, s_name) =>
        if(cache.styleHierarchy.contains(s_name.trim))
            s_name.trim :: styles else styles)
```

Listing 4.23: Auszug aus Code sammeln der Elterninstanzen

Dafür wird in der vom Parser mitgegebenen ClassHierarchy (*styleHierarchy*) nach den Namen der Eltern gesucht (wieso hierbei die ClassHierarchy Instanz nicht angesprochen werden muss und warum der Liste aus Styles augenscheinlich immer nur Strings zugewiesen werden, wird in 4.5.1 erklärt). Anschließend werden die verschiedenen Variablen, die später an den Konstruktor übergeben werden vordefiniert (siehe Listing 4.24).

```
var description:Option[String]=relevant{_.description}
var line_color:Option[Color]=relevant{_.line_color}
var line_style:Option[LineStyle]=relevant{_.line_style}
var line_width:Option[Int]=relevant{_.line_width}
```

Listing 4.24: Auszug aus Code Zuweisung der Felder der Elternelemente

Auffällig ist hierbei, dass die Zuweisungen beinahe auch für nicht Programmierer lesbar sind. Wie zum Beispiel die erste Zuweisung in Listing 4.24: “*Weise der description die relevante bestehende description zu*“. Hinter diesem Aufruf

verbirgt sich eine Funktion, die in `ClassHierarchy` definiert wurde, um einfach auf die Attribute der Elternklasse zugreifen zu können (siehe Listing 4.25).

```
def mostRelevant[T, C](stack: List[C])(getter: C => Option[T]): Option[T] = {  
  for (parent <- stack) {  
    if (getter(parent).isDefined)  
      return f(parent)  
  }  
  None  
}
```

Listing 4.25: Auszug aus Code Funktion `mostRelevant`

Die Methode `mostRelevant` macht Gebrauch von mehreren nützlichen Eigenschaften von Scala. Zunächst wird die Methode mit Typparametern beschrieben. Was diese jeweils beschreiben sollen wird gleich klar. Sinn und Zweck der `mostRelevant` Methode ist es dem Benutzer zu ermöglichen, einen generischen `getter` auf beliebige Felder der Oberklassen zu bieten. Sowohl `Style` als auch `Shape` führen eine Liste mit den Instanzen, die sie erweitern. Genauer, führen `Style` und `Shape` eine Art *Stack*, denn bei der Vererbung soll hier das *latest-Bound-Principle* gelten. Dieser Stack wird als erster Parameter übergeben. Als nächstes sieht man eine weitere Parameterliste, die eine Funktion erwartet. Da die Funktion komplett typunabhängig funktionieren soll, wurden Typparameter benutzt. `C` stellt in diesem Fall den Typ der Modellklasse dar, `T` hingegen ist der Typ des gewünschten Elternattributs. Die Funktion bildet eine Modellklasse (`C`) auf den gewünschten Typ (`T`) ab und ist somit als *Getter* Methode zu verstehen. `MostRelevant` iteriert also über die Elternklassen, wobei die Relevanz beziehungsweise die Priorität vom ersten Element der Liste, bis zum Letzten absteigend ist. Da alle Felder der Modellklassen *Options* sind, kann nun für jede Superklasse geprüft werden, ob das gewünschte Feld definiert ist. Dieses gewünschte Feld wird wiederum über die mitgelieferte Funktion ermittelt. Wird in Keiner der Elternklassen ein passendes Feld gefunden, wird `None` zurückgegeben. Da der zweite Parameter in einer eigenen Parameterliste angegeben ist, kann das entsprechende Argument anstatt in runden, in geschweiften Klammern angegeben werden und fügt sich somit wie eine neue Kontrollstruktur in den Code ein. Momentan müsste ein Aufruf aus dem `Style Companion` Objekt wie in Listing 4.26 aussehen.

```
var description: Option[String] =  
  ClassHierarchy.mostRelevant(extendedStyles){ _.  
    description }
```

Listing 4.26: Beispielhafter Aufruf von `mostRelevant`

Um selbst diesen Aufruf noch zu verkürzen wird sich einfach einer weiteren Hilfsmethode, namens *relevant* bedient (siehe Listing 4.27).

```
def relevant[T](f: Style => Option[T]) =
  ClassHierarchy.mostRelevant(extendedStyle) {f}
```

Listing 4.27: Auszug aus Code Hilfsfunktion relevant

Diese setzt die Elternliste schon einmal voraus und erwartet ab hier nur noch eine Funktion, welche aus einem Style Element ein beliebiges T extrahiert. Da der Scala Compiler über die mitgelieferte Funktion den Typparameter für die relevant Methode ermitteln kann, wird sogar der Typ der zuzuweisenden Variable inferiert. Somit reduziert sich das Abrufen des am meisten relevanten Feldes eines beliebigen Typs einer Elternklasse erheblich (siehe Listing 4.28).

```
var description = relevant{ _.description }
```

Listing 4.28: Beispielaufruf um die latest-Bound description der Elterninstanzen zu ermitteln

Da Scala Funktionen als *First-Class Objects* behandelt werden, können sie eben auch als Argumente vergeben werden (vgl. Esser 2011, S.244). Methoden werden hierfür implizit in Funktionen umgewandelt. In Sprachen, in denen Funktionen anders behandelt werden, wäre dies so nicht möglich. Um bei der Stylegenerierung den Wert einer Elternklasse für das gewünschte Feld zu ermitteln, müsste für jedes einzelne Attribut eine eigene Funktion erstellt werden. Diese iteriert durch die Elternklassen und führt dann eine nicht variable eindeutige *Getter* Methode aus. Scala kommt hier mit einer einzigen Funktion und einer weiteren Hilfsfunktion (relevant), welche nur dem Komfort dient, aus.

### 4.3.1 Vererbung bei Shapes

Die Vererbungslogik hat sich bisher allgemein auf alle Felder der Modellklassen bezogen. Im Falle einer erweiterten Shape, sind zunächst zwei Fragen prinzipiell zu klären:

- Werden geerbte geometrische Figuren referenziert, oder müssen tiefe Kopien davon erstellt werden?
- Werden geometrische Figuren an geometrische Figuren oder an Shapes vererbt?

#### Referenzierung oder tiefe Kopie geerbter Felder?

Eine Referenzierung der geerbten geometrischen Figuren ist wünschenswert, da insbesondere große geschachtelte Bäume aus geometrischen Figuren sehr viel speicherschonender behandelt werden würden. Gefahrlos umsetzen lässt sich dies aber nur wenn garantiert wird, dass eben jene geometrischen Figuren immutabel sind. Der Ansatz ist, alle Shape Member über ein *val*, also konstant zu definieren. Dies setzt jedoch voraus, dass die Anforderung einer späteren Änderung der Felder trotz der Shape Vererbung nicht nötig ist. Daraus ergibt sich die nächste Frage.



## Geerbte geometrische Figuren

Wird eine Shape A definiert, die ein Rechteck enthält und wird nun eine Shape B definiert, die selber eine Ellipse enthält und zusätzlich Shape A erweitert, so ergeben sich nun zwei Möglichkeiten, die Shapevererbung umzusetzen (siehe Listing 4.30 und 4.29).

```
shape B extends A {  
    ellipse {  
        rectangle { ... }  
    }  
}
```

Listing 4.29: Erweiterung der geometrischen Figuren anhand z.B. der Tiefe der Elemente

```
shape B extends A {  
    ellipse { ... }  
    rectangle { ... }  
}
```

Listing 4.30: Einfügen separater Bäume

Würde die Shapevererbung bedeuten, dass die bereits bestehenden geometrischen Figuren z.B. ausgehend von der Tiefe erweitert werden (Listing 4.29), würde dies ebenso bedeuten, dass geometrische Figuren auch nach ihrer Erzeugung verändert werden können müssen. Damit wären sie nicht immutabel und könnten nicht gefahrlos referenziert werden. Ergo wäre die Speicherauslastung höher. Die Lösung in Listing 4.30 schlägt vor, vererbte geometrische Figuren als separaten Baum in die neu erzeugte Shape einzufügen. Die im Beispiel beschriebene Ellipse wird dabei nicht nachträglich verändert und kann somit als immutabel definiert werden. Nun reicht eine einfache Referenz auf die vererbten Felder und senkt die Speicherauslastung.

### 4.3.2 Transitivität der Styleeigenschaften

Eine wichtige Anforderung an die Objektstruktur ist, dass Styleinformationen in Shapes an die beinhalteten geometrischen Figuren weitergegeben werden. So wird erreicht, dass geometrische Figuren mit gleichen Farb-, Font- und anderen Eigenschaften erstellt werden. Wenn also einer Shape A ein Style S zugewiesen wird und ferner, A einen Baum aus geometrischen Figuren enthält, muss gewährleistet werden, dass alle unterliegenden geometrischen Figuren von A über die selben Styleinformationen von S verfügen, wie ihr Eltern Shape. Dabei hört es noch nicht auf, denn sollte nun eine der geometrischen Figuren ebenfalls einen Style (S1) zugeordnet bekommen, so müssen nun dessen unterliegende geometrische Figuren über die Styleinformationen von S und S1 verfügen. Dabei hört es ebenfalls noch nicht auf. Denn entgegen einer Stylereferenz, ist es geometrischen Figuren außerdem auch möglich einen anonymen Style innerhalb

ihres Scopes zu definieren. Durch diesen anonymen Style, werden die vorherigen Styleinformationen jedoch **nicht** revidiert, sondern wenn überhaupt erweitert. Es ist also zu beachten, dass Styleinformationen in die Tiefe weitergegeben werden und dabei von tieferen Styles, ob anonym oder nicht, gegebenenfalls erweitert werden. Ein vereinfachtes Beispiel demonstriert anhand zweier vordefinierter Styles und einer Shape mit mehreren geometrischen Figuren was gemeint ist (siehe Listing 4.31).

```
style S1 {
    color = blue
}

style S2 {
    transparency = 0.5
}

shape A style S1 {
    rectangle {
        ellipse {
            rectangle style S2 {
                style {
                    color = red
                }
            }
        }
    }
}
```

Listing 4.31: Beispiel zur Verdeutlichung transitiver Styleeigenschaften

Diese Definitionen sieht fertig umgesetzt aus wie in Abbildung 4.4.

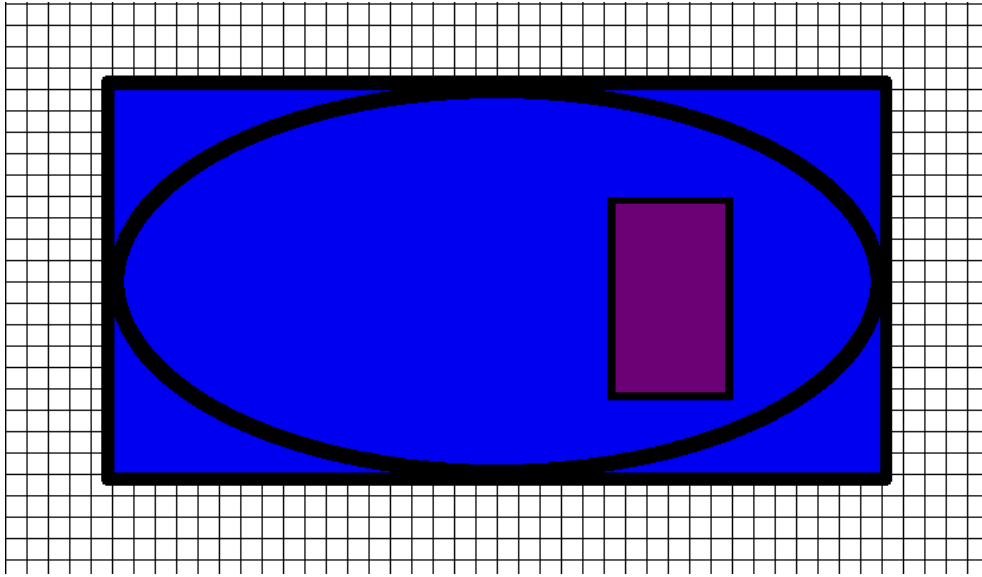


Abbildung 4.4: Beispiel transitiver Styleinformation

In 4.4 ist zu sehen, dass die Ellipse die Farbeigenschaften des äußeren Rechtecks übernimmt, das innere Rechteck jedoch sowohl seine eigene Farbeigenschaft (rot) definiert und außerdem die Transparenz aus *S2* umgesetzt ist. Bei der Vererbung wurde bisher wie bereits erwähnt darauf geachtet, Felder der Superklassen direkt bei der Erzeugung der neuen Instanzen weiterzugeben, um rekursives Suchen der entsprechenden Attribute in Bäumen zu vermeiden. Ebenso sind hier die Prinzipien der transitiven Style-Eigenschaften über Vererbung gelöst und auch in diesem Fall wird darauf verzichtet nur Referenzen zu speichern, über die Elternknoten erreicht werden können. Die Parserregel für ein *GeometricModel* beziehungsweise dessen "Skizze" dem *GeoModel* erkennt anonyme Styles während dem Parsen. Da anonyme Styles **keine** Informationen über Elternpaare besitzen müssen, werden sie direkt instantiiert und in einem Cache hinterlegt. Der Attributliste des entsprechenden *GeoModel* wird lediglich der Name des anonymen Styles mitgegeben (tatsächlich bekommt der anonyme Style intern einen generierten Namen, den der Benutzer nicht kennt). So kann weiterhin das einheitliche Medium einer Argumentenliste aus String Tupeln an die Factorys weitergegeben werden. Die entsprechende Factory löst das Argument schlussendlich über den Namen auf und erhält wieder die eigentliche Styleinstanz. Styles werden wie andere Attribute auch in den entsprechenden Parserklassen aufgelöst und an die Konstruktoren der Modellklassen weitergegeben. Am Beispiel des *CommonLayouts* wird deutlich wie die Style Vererbung funktioniert. Wie die anderen Parser auch, definiert das *CommonLayout* zuerst Variablen, welche ausgehend von geparster und geerbter Information initialisiert und später an den Konstruktor überreicht werden. Für den Styleparameter erzeugt das *CommonLayout* also zunächst einen neuen Style, der sich aus dem Style der Elterninstanz und dem eigenen zusammen setzt (siehe Listing 4.32).

```
var styl:Option[Style] = Style.generateChildStyle(cache,
  parentStyle, geoModel.style)
```

Listing 4.32: Erzeugen eines neuen Styles über andere Styleinstanzen

Hier ist zu sehen wie die Styleinformationen der nächst höheren Instanz und der eigenen Style Referenz ausgewertet werden. Mehr zu der Funktion *generateChildStyle* im weiteren Verlauf. Wird unter den Attributen nun beim Pattern Matching der Name eines anonymen Styles gefunden, wird dieser ebenfalls in die bestehenden Styleinformationen eingearbeitet (siehe Listing 4.33).

```
case anonymousStyle:String if cache.styleHierarchy.
  contains(anonymousStyle) =>
  styl = Style.generateChildStyle(cache, styl, Some
    (anonymousStyle))
```

Listing 4.33: Auszug aus Code Erzeugen von Kindstyles

Was genau macht nun *generateChildStyle*? *generateChildStyle* ist eine Funktion des Style Companion Objekts (siehe Listing 4.34).

```
def generateChildStyle(cache: Cache, parents:Option[Style
  ]*):Option[Style] ={
  val parentStyles = parents.filter(_.isDefined)
  if(parentStyles.length == 1) return parentStyles.head
  else if(parentStyles.isEmpty) return None
  val childName =
    "(child_of -> "+parentStyles.map( p => p.get.name
      +{if(p != parentStyles.last)" & "else ""}).
      mkString+")"
  Some(Style(childName, Some(parentStyles.toList.map(i
    => i.get.name)), List[(String, String)](), cache))
}
```

Listing 4.34: Auszug aus Code Funktion generateChildStyle

Die Funktion erwartet eine variable Anzahl an Styles (Scala erlaubt diese *varargs* durch die Kennzeichnung \*), prüft anschließend, ob es sich um gültige Argumente handelt. Im Falle, dass mehr als ein gültiges Argument unter den Eltern ist, wird ein **neuer** Style über die Factory erzeugt, der die entsprechenden Elternteile mitbekommt. Die Style Factory kümmert sich nun wieder darum, die Felder des neuen Styles anhand des *latest Bound Prinzips* zu erben, wodurch ein neuer “Kind“ Style entsteht. Im Folgenden wird nun ein weiteres Problem adressiert.

### 4.3.3 Modell und Skizze; Problematik der Instantiierungsreihenfolge

Über Style, Connection und Shape ist es möglich in der Definition des allumfassenden Diagrams ebenfalls eine Style Referenz anzugeben. Effektiv soll

hierdurch ein *Corporate Design* realisiert werden können, welches sämtliche Eigenschaften an unterliegende Connections, Shapes und somit auch geometrische Figuren überträgt, Gerhart and Boger (**year?**), S.3. Um wie bereits beschrieben die Möglichkeit zu bieten, die Modellklassen in Zukunft parallelisiert bearbeiten zu können, wurden alle Attribute der Modellklassen als *vals* gekennzeichnet und sind somit Konstant. So können ruhigen Gewissens funktionale Programmierparadigmen darauf angewendet werden und entsprechender Code kann hoch parallel arbeiten. Entsprechend ist es nicht möglich eine (beispielsweise) Shape zu erzeugen, anschließend ein Diagramm zu erstellen, das auf die bereits fertige Shape referenziert und ihr somit nachträglich Styleinformationen übertragen zu wollen. Die Shape ist ja bereits erstellt und alle Felder, inklusive der Styleinformationen, sind bereits aufgelöst. Hier wird nun die selbe Taktik angewandt, wie zuvor bei den geometrischen Figuren 4.4. Da die Erzeugung der Shape Instanzen folglich erst erfolgen darf, wenn mögliche Corporate Styleinformationen bereits gegeben sind, werden Shapes zunächst über eine Methode namens *shapeSketch* geparsed. *ShapeSketch* sammelt für eine angehende Shape, wie das GeoModel für angehende GeometricModels alle geparsten Informationen, löst diese aber noch **nicht** auf. Beim Einlesen eines Diagrams, werden sowohl Style als auch Shape referenziert. Erst also wenn alle nötigen Informationen vorhanden sind, wird während des Parsing Prozesses des Diagrams die Shape Skizze(*ShapeSketch*) mit dem Corporate Style zu der eigentlichen Shape aufgelöst. Hierfür verfügt die ShapeSketch Klasse über eine Methode namens *toShape*, die die gespeicherten Attribute an die Shape Factory übergibt. So erzeugte “Skizzen“ sind somit nur Container um angehende Modellklassen zu einem beliebigen Zeitpunkt auflösen zu können (siehe Listing 4.35).

```
case class ShapeSketch(name:String,
                      parents:Option[List[String]],
                      style:Option[Style],
                      attrs:List[(String, String)],
                      geos:List[GeoModel],
                      descr:Option[(String, String)],
                      anch:Option[String],
                      cache: Cache){
  def toShape(corporateStyle:Option[Style]) =
    Shape(name, parents, Style.generateChildStyle(cache
      , corporateStyle, style),
      attrs, geos, descr, anch, cache)
}
```

Listing 4.35: Auszug aus Code Definition der ShapeSketch container Klasse

Außerdem wird der nun vorhandene Corporate Style und der Style der Shape, über die bereits vorgestellte Funktion *Style.generateChildStyle* kombiniert. Da der Corporate Style auf diese Weise ganz oben in der Shape eingesetzt wird, verteilt er sich wie bereits beschrieben (siehe 4.3.2) automatisch auf die unterliegenden geometrischen Formen, da die entsprechenden Factorys die über-

liegenden Instanzen analysiert und entsprechend erweitert. Das Problem der Erzeugungsreihenfolge zieht sich durch das gesamte Projekt. Die Abhängigkeit zu Style Instanzen lässt es nicht zu, irgendeine Modellklasse vor dem Diagramm zu erzeugen, da Styleinformationen bis hin zum Diagramm immer erweitert werden können. Folglich ist es zwingend erforderlich für sämtliche Modellklassen, die von Style abhängen, Container zu erstellen, die die unaufgelösten Attribute der Modellklasse beinhalten und es ermöglichen die Erzeugung der eigentlichen Instanz auf einen späteren Zeitpunkt zu verschieben. So findet sich zu jeder Modellklasse und jeder Hilfsklasse, die von Style abhängt, eine entsprechende Containerklasse, die mit *Sketch* gekennzeichnet ist:

**Diagram** Node und NodeSketch; Edge und EdgeSketch

**Shape** Shape und ShapeSketch; GeometricModel und GeoModel(einzige Ausnahme, endet nicht mit Sketch)

**Connection** Connection und ConnectionSketch

Diese Lösung resultiert aus der Vorgabe, dass alle Felder der Modellklassen immutabel sein sollen und dem gewählten Lösungsweg, vererbte Eigenschaften bei der Objekterzeugung zu übergeben und nicht über Referenzen auf Elternelemente beim Aufrufzeitpunkt zu suchen. Der so gewonnene Vorteil während der Benutzung der Modelle mit besseren *Response Times* rechnen zu können, da die Rekursivität des Problems aufgelöst wurde, wird also im Endeffekt durch einen zusätzlichen Overhead an Speicherauslastung wieder wett gemacht. Erzeugte Container/Skizzen können nach Auflösung zur entsprechenden Instanz der Modellklasse nämlich nicht verworfen und dem *Garbage Collector* überlassen werden, da weitere Diagrams ebenfalls noch in der Lage sein müssen unberührte Skizzen der Modellklassen benutzen zu können. Die Klasse *SprayParser* stellt die eigentliche Schnittstelle zum Parser dar. Sie enthält alle notwendigen Methoden um das Parsen eines Strings einzuleiten.

**parseStyle** Zum parsen von n Styles

**parseShape** Zum parsen von n Shapes

**parseAbstractShape** Zum parsen von n abstrakten Shapes

**parseConnection** Zum parsen von n Connections

**parseAbstractConnection** Zum parsen von n abstrakten Connections in einem String

**parseDiagram** Zum parsen von n Diagrams in einem String

Diverse weitere Methoden, sind gekapselt um den SprayParser nach außen hin möglichst simpel und benutzerfreundlich zu gestalten. Wird eine Methode mit der Kennzeichnung *Abstract* zum Parsen einer Shape oder Connection benutzt, werden tatsächliche Shape bzw. Connection Instanzen zurückgeliefert, jedoch auch für diesen Prozess die entsprechenden Container Skizzen angelegt. Das dient dem Zweck, dass es für Diagrams auch möglich ist abstrakte Shapes zu

referenzieren hat aber eine sehr wichtige Konsequenz. Referenziert ein Diagramm eine abstrakte Shape (oder Connection), wird die entsprechende *ShapeSketch* verwendet, damit corporate Styles beachtet werden können. Die bereits bestehende Shape Instanz wird im Cache sodurch aber überschrieben, da der entsprechende Container neu umgewandelt wird und der bestehende Eintrag im Cache mit dem entsprechenden *Key* (Name der Shape) überschrieben wird. Abstrakte Shapes sollten also nur dann von Diagramms referenziert werden, wenn absichtlich darauf verzichtet wird, dass Unterklassen der abstrakten Shape den corporate Style des Diagramms dadurch **nicht** implizit übernehmen. Deren Styleinformationen sind ja bereits aufgelöst und können wie bereits erklärt nicht nachträglich aktualisiert werden. Um dieses Prinzip zu verdeutlichen ein Beispiel (siehe Listing 4.36). Es wird eine Shape *A* definiert, welche einen bestimmten Style benutzt und als abstrakte Shape geparst wird. Eine weitere Shape *B*, die *A* erweitert, wird definiert und als normale Shape geparst. Wird nun ein Diagramm definiert, das unbedingt die abstrakte Shape *A* referenzieren soll und ein corporate Style namens *CorporateStyle* benutzt und anschließend eine neue Shape *C* definiert, welche die abstrakte Shape *A* erweitern soll, dann hat nun auch die Shape *C* die Styleinformationen von sowohl dem *CorporateStyle* und dem *DefaultStyle*, da die Shape, die sie referenziert von der Diagramm Definition neu erzeugt worden ist.

```
val abstractShape = """shape A style DefaultStyle {...}"""
parser.parseAbstractShape(abstractShape)

val normalShape = """shape B extends A {...}"""
parser.parseShape(normalShape)

val diagram = """diagram D ... (style:CorporateStyle){
    node ... {
        shape:A(...)
    }
}"""
parser.parseDiagram(diagram)

val normalShape2 = """shape C extends A {...}"""
parser.parseShape(normalShape2)
```

Listing 4.36: Beispiel Aufrufe zum Provozieren einer möglicherweise ungewollten Situation durch Vererbung

Da abstrakte Shapes in der Regel nur erzeugt werden um gemeinsame Attribute anderer Shapes zu vereinheitlichen oder um tests mit Shapes durchzuführen, ohne ein Diagramm erstellen zu müssen ist der beschriebene Fall normalerweise nicht von Belang, da Diagramms normalerweise auf keine abstrakte Shape verweisen sollten. Hier muss überlegt werden, ob sich aus dem Referenzieren von abstrakten Shapes ein zusätzlicher nützlicher Effekt ergeben kann oder ob dadurch eher eine potenzielle Fehlerquelle eingeführt wird. Je nach dem muss über eine

Prüfung verhindert werden, dass Shapeinstanzen überschrieben werden können. Wird eine Shape bzw. Connection über *parseShape* bzw. *parseConnection*, also ohne die Abstract Kennzeichnung, geparkt werden zunächst **keine** tatsächlichen Instanzen der Modellklassen erzeugt, sondern nur die der Container Klassen *ShapeSketch* bzw. *ConnectionSketch*. Jetzt erscheint es an dieser Stelle paradox, dass die Methoden mit der *Abstract* Kennzeichnung tatsächliche Instanzen der Modellklassen erzeugen und die ohne *Abstract* Kennzeichnung nur Container Klassen. Aus Anwendersicht ergibt dies schon mehr Sinn. Als Anwender definiert man Shapes und Connections nämlich eigentlich nur dann, wenn man sie später auch in einem Diagram referenziert. Mit anderen Worten sind die tatsächlich benutzten Shapes und Connections die, die über die Methoden *parseShape* bzw. *parseConnection* eingelesen werden. Shapes und Connections, die man nur dafür einlesen will, um ähnliche Eigenschaften zusammenzufassen und durch Unterklassen zu diversifizieren (sinngemäß also abstrakt), werden über die Methoden *parseAbstractShape* bzw. *parseAbstractConnection* eingelesen.

### Kritik an gewählter Taktik

Im Falle, dass mit Referenzen auf Styles der Elternobjekte gearbeitet werden würde, wäre es möglich die Styleinformationen der Elterninstanz und die eigenen erst zu erzeugen, wenn sie benötigt werden. So bestünde kein Bedarf für die vorgestellten Container Klassen. Je nach dem wie tief die Vererbungshierarchie ist, würde dies jedoch beträchtliche Nachteile im Bezug auf die *Responsetime* des später erzeugten graphischen Editors mit sich bringen. Die Skizzen Klassen zu den jeweiligen Modellklassen wurden in der *SprayParser.scala* Datei definiert, da andere Parsingstrategien eventuell nicht mehr abhängig von ihnen sind.

### 4.3.4 Generatorenklassen

Die Generatorenklassen greifen auf Informationen der Modellklassen zurück um sie auf ein gewünschtes Format abbilden zu können. Jede Modellklasse hat ein entsprechendes Pendant als Generator, so hat zum Beispiel Shape einen *ShapeGenerator*. Da in den Generatoren viele Informationen verarbeitet werden, sind diese wiederum in mehrere Generatorenklassen aufgeteilt. ShapeGenerator wird unter anderem durch einen *GeneratorShapeDefinition* und *GeneratorInspectorDefinition* unterstützt. So werden auch hier Verantwortungen logisch aufgeteilt. Prinzipiell konnten die Generatorenklassen größtenteils von der Xtext Vorlage übernommen werden. Konzeptuell wurde so nichts verändert, außer dem zugriff auf die Modellklassen und der Formatierung des auszugebenden Javascript Codes. Scala bietet zum dynamischen Erstellen von Strings Stringinterpolation zur verfügung. So kann (wie man es beispielsweise aus Shell Skripten kennt) über ein "\$" Zeichen das Ausführen von Codeblöcken eingeleitet werden. Dadurch ist es möglich sogar innerhalb eines Strings dynamisch auf Kontrollstrukturen wie *for Schleifen* (und prinzipiell auf die ganze Macht Scalas) zurückzugreifen.



## 4.4 Ausführliches Beispiel

Nun folgt ein ausführliches Beispiel anhand einer Shapedefinition (siehe Listing 4.37), zur Verdeutlichung der Schritte, die beim Parsen einer Modellklasse, durchlaufen werden.

```
shape EClassShape style B{
  size-min (width=4, height=6)
  rectangle {
    style (line-width=2)
    position (x=2, y=0)
    size (width=10, height=3)
    ellipse {
      position (x=0, y=36)
      size (width=30, height=30)
    }
  }
}
```

Listing 4.37: Beispielformatdefinition für eine Shape

Soll eine Shapedefinition wie 4.37 eingelesen und direkt in eine Shape Instanz umgewandelt werden, wird sie zunächst im SprayParser (der Standard Parser-Klasse) durch folgende Regeln/Parser überprüft:

```
[ 1] private def shapeSketch:Parser[Shape] =
[ 2]   ("shape" ~> ident) ~
[ 3]   (("extends" ~> rep(("(!style)".r ~> ident)<~ ",?"
[ 4]   ".r"))?) ~
[ 5]   (("style" ~> ident)?) ~
[ 6]   ("{" ~> rep(shapeAttribute)) ~
[ 7]   rep(geoModel) ~
[ 8]   (descriptionAttribute?) ~
[ 9]   (anchorAttribute?) <~ "}" ^^
[10]   {
[11]     case name ~ parent ~ style ~ attrs ~ geos ~
[12]       desc ~ anch =>
[13]       ShapeSketch(name, parent, style, attrs, geos,
[14]         desc, anch, cache)
```

Listing 4.38: Auszug aus Code Methode zum Parsen einer Shape(Sketch)

Warum die Methode/Regel in Zeile eins *shapeSketch* heißt wurde in Abschnitt 4.3.3 erklärt. Zeile zwei filtert den identifier der Shape. Zeile drei prüft auf eine optionale Angabe erweiterter Shapes. Hierbei dürfen diese nicht „style“ heißen, da die Regel sonst verwirrt wäre, erwartet sie doch in Zeile vier eine optionale Angabe für einen benutzten Style. Zeile fünf sucht nun *n* mal nach einer Regel namens *shapeAttribute* (siehe Listing 4.39)

```
private def shapeAttribute = shapeVariable ~ arguments ^^
  {case v ~ a => (v, a)}
private def shapeVariable = ("{" + Shape.
  validShapeVariables.map(_+"|").mkString+"}") .r ^^
  {_.toString}
```

Listing 4.39: Auszug aus Code Regeln zum Parsen der Felder

*validShapeVariables* ist eine Sammlung von Strings, welche die bekannten Felder einer Shape darstellt. Diese Sammlung wird zu einem regulären Ausdruck umgewandelt und unter der Regel *shapeVariable* benutzt. Die Regel *arguments* entstammt dem Trait *CommonParserMethods*, welches jedem Parser zur Verfügung steht. So kann *shapeAttribute* Tupel von Attribut & Wert Paaren bilden. Shape Attribute sind lediglich primitive Datentypen und daher nicht interessant weiterzuverfolgen. In Zeile sechs von Listing 4.38 wird ebenfalls *n* mal die Regel *geoModel* angewandt, welche die einzelnen Regeln einer geometrischen Figur abbildet und ein Objekt zurückliefert, das als Container(4.3.3) fungiert (siehe Listing 4.40).

```
private def geoModel: Parser[GeoModel] =
  geoIdentifier ~
  (((("style" ~> ident)?) <~ "{") ~
  rep(geoAttribute|anonymousStyle) ~
  (rep(geoModel) <~ "}")) ^^
  {
    case name ~ style ~ attr ~ children =>
      GeoModel(name, {if(style.isDefined) Some(style.
        get) else None }, attr, children, [...])
  }
```

Listing 4.40: Auszug aus Code Regel zum Parsen geometrischer Figuren

Hierbei werden nicht direkt Instanzen der eigentlich gewünschten *GeometricModel* Klasse erzeugt, da es für die geometrischen Figuren zwingend erforderlich ist einerseits ihre entsprechende Eltern Shape Instanz zu kennen. Diese existiert zu diesem Zeitpunkt jedoch noch nicht. Außerdem müssen sie ihre Eltern *GeometricModel* Instanzen kennen, welche ebenfalls erst erzeugt werden können, wenn ihre Kinder existieren. Alle Felder sollen Konstanten sein und nicht nachträglich gesetzt werden. Diese Teufelsspirale wird gelöst, indem zunächst Container Objekte erstellt werden, die die Informationen speichern und umgewandelt werden, sobald ihre Abhängigkeiten geklärt sind. Sind alle Regeln überprüft und keine Fehleingaben oder fehlende Eingaben identifiziert worden, gibt es von nun an zwei Möglichkeiten weiter zu verfahren. Einer der Gründe, warum Container erstellt werden müssen wird in 4.3.2 erklärt. Wird der Container jetzt bereits der Fabrik für Shapes übergeben und somit eine fertige Shape Instanz erzeugt, können der Shape Instanz nachträglich **keine** weiteren Informationen mehr gegeben werden. Wird die Shape bereits jetzt erzeugt, existiert sie zwar als valide Shape, ist aber nun mehr als abstrakte Shape zu betrachten, da neue Shapes sie

zwar erweitern können, sie aber nicht selber in einer Diagram Instanz referenziert wird. Das Diagram würde versuchen nachträglich Styleinformationen in die Shape einzuarbeiten. Hierzu mehr in Abschnitt 4.3.3. Um die Bearbeitungsschritte einer Shape zu verdeutlichen reicht es hier, den Container, welcher die Information der angehenden Shape enthält, nun direkt an die Shape Fabrik (das *Companion Object* der Shape Klasse) zu übergeben. Wie man in Zeile 11 sieht, wird beim Parsen einer Shape der Container ShapeSketch zurückgeliefert. Um entweder zu Testzwecken, oder um eine abstrakte Shape zu erzeugen, kann ein String mit Shapedefinition deshalb über die Methode *abstractShape* geparsed werden (siehe Listing 4.41).

```
private def abstractShape:Parser[Shape] = shapeSketch ^^
  {case sketch => sketch.toShape(None)}
```

Listing 4.41: Auszug aus Code Methode zum direkten Parsen von Shapes

Wie man sieht greift *abstractShape* auf die Regel/Methode *shapeSketch* zurück, erzeugt jedoch anschließend direkt über *ShapeSketch*'s Methode *toShape*, welche lediglich ein Aufruf an die *Shapefactory* weiterleitet, eine valide Shape. Ab dem Aufruf *toShape* übernimmt der spezialisierte Shape Parser im *Companion Objekt* der Shape Klasse das Ruder und bereitet die erhaltenen Argumente so auf, dass diese an den eigentlichen Shapekonstruktor weitergegeben werden können. Dabei werden zunächst die Superklassen aufgelöst und erweitert (siehe Abschnitt 4.3). Anschließend werden über ein Pattern Matching (siehe Listing 4.42) die erhaltenen Attribut- und Wert Tupel aufgelöst. Beispielsweise das "size-min" Attribut, das zuvor in Listing 4.37 gegeben wurde.

```
attributes.foreach{
  case ("size-min", x) =>
    val opt = parse(width_height, x).get
    if(opt.isDefined){
      size_width_min = Some(opt.get._1)
      size_height_min = Some(opt.get._2)
    }
  ...
}
```

Listing 4.42: Auszug aus Code Iterieren durch Attribute mit Pattern Matching

Auch hierbei wird an weitere Regeln delegiert (z.B. *width\_height*). Herausgefilterte Werte werden vordefinierten Variablen zugewiesen, welche später an den eigentlichen Konstruktor übergeben werden. Sind so alle Attribute der Shape aufgelöst wird der private Konstruktor der Shape companion Klasse aufgerufen. Es ist darauf zu achten, dass alle Felder der erweiterbaren Modellklasse Konstanten sind (siehe Abschnitt 4.3). Dazu müssen dem Shapekonstruktor die "Skizzen" (die Container) der geometrischen Figuren übergeben werden. GeometricModels sollen ebenfalls immutabel sein und müssen daher bereits bei ihrer Erzeugung ihre Eltern Shape kennen. Würden die geometrischen Figuren vor der Shape Instantiierung erzeugt werden, müsste die Referenz auf die Eltern

Shape nachgereicht und die entsprechende Referenz als *var* definiert werden. Da die Erzeugung der GeometricModels im Konstruktor erfolgt, kann die umschließende Shape über *this* referenziert werden. Von nun an befindet man sich im privaten Konstruktor der Shape Modellklasse. Die wohl wichtigste Aufgabe des Shape Konstruktors ist es die geometrischen Figuren zu erzeugen. Hierbei muss der Konstruktor jedoch auch beachten, dass von einer möglichen Elternshape weitere geometrische Figuren geerbt werden können (siehe Abschnitt 4.3.1). Deshalb werden die neu geparsten und die geerbten geometrischen Figuren zusammen in eine Liste *gemerged*. Die hier benutzte Methode *parseGeometricModels* (siehe Listing 4.43), iteriert durch die Container der geometrischen Figuren und erzeugt vollwertige GeometricModels.

```
private def parseGeometricModels
(geoModels:List[GeoModel], parentStyle:Style) =
  geoModels.foldLeft(List[GeometricModel]())((list, g)
    => {
      val newGeo = g.parse(None, parentStyle)
      if (newGeo.isDefined) newGeo.get :: list else
        list
    })
```

Listing 4.43: Auszug aus Code Methode parseGeometricModels

In Listing 4.43 wird die *parse* Methode der Container aufgerufen, welche anhand eines Matchings entscheidet um welche geometrische Figur es sich explizit handelt (z.B. Ellipse oder Rechteck etc.). Anhand dieses Matchings wird dann der spezialisierte Parser der vielen verschiedenen Klassen aufgerufen, die eine geometrische Figur beschreiben. Im Falle der obigen Beispielshape wurde ein Rechteck definiert, welches eine Ellipse beinhaltet. Ab hier geht es also in der *Rectangle Fabrik* weiter (dem Companion Objekt der Rectangle Klasse). Hier werden wie in allen anderen Parsern die mitgegebenen Attribute über ein Matching aufgelöst, wobei hierfür wieder eventuell auf weitere spezialisiertere Parser zurückgegriffen wird. Zum Beispiel wird das Rectangle seine Attribute an die Parser für das *CommonLayout* und die *CompartmentInfo* weitergeben. Erst wenn deren Fabrikmethoden valide Ergebnisse zurück liefern wird das Rechteck erzeugt. Der private Konstruktor der Rectanlge Klasse bekommt wiederum den Container der im Beispiel beschriebenen Ellipse mit und so geht der Vorgang in der Ellipse Fabrik von vorne los.

## 4.5 Cache

Bisher wurde bereits einige male erwähnt, dass sich Modellklassen untereinander referenzieren. Diese Referenzierungen können aber Aufgrund der Problematik der Instantiierungsreihenfolge nicht während des Parsingvorgangs erfolgen. Sonst würde auf Objekte verwiesen werden, die selber bereits erst in der Mache sind. Oder es geht schlicht um die Problematik, dass Styles und Shapes unabhängig voneinander definiert werden können, Shapes jedoch namentlich auf

Styles verweisen. Selbst wenn es also nur darum geht Referenzen auf erstellte Styles zu behalten um sie nicht an den *Garbage Collector* zu verlieren, ist ein Mechanismus nötig, der sich entsprechende Referenzen behält und wiederum bereitstellt. Da sich Diagrams, Shapes und Styles untereinander referenzieren können sollen, muss der Parser also in der Lage sein Instanzen zu erzeugen und gleichzeitig auch zu verwalten. Hierbei wurde stets darauf geachtet, dass der SprayParser keine statischen Zustände hält. Ausgehend davon, dass der Parser zukünftig in einer Multi-Client-fähigen Umgebung eingesetzt wird, ist so garantiert, dass sich verschiedene Clients niemals die selben Ressourcen teilen und somit eventuell inkonsistente Daten produzieren. Schon eine gemeinsame Abhängigkeit eines Namensraums wäre nicht wünschenswert, da Modellklassen mit gleichen Namen sonst je nach Implementierung entweder Exceptions auslösen würden oder sich gegenseitig überschreiben könnten. Wenn aber keine statischen Zustände gehalten werden kann jeder Client mit einer eigenen SprayParser Instanz kommunizieren. Außerdem kann so auf komplizierte Synchronisationsmechanismen verzichtet werden. Es wird beispielsweise zunächst eine Shape geparsed, auf die ein Diagramm verweisen soll. Jedoch kann das Diagramm die Shape nicht referenzieren, wenn es die entsprechende Referenz gar nicht kennt. Um die Referenz auf die Shape also nicht zu verlieren, muss sie lokal gespeichert und über eine Eigenschaft erreichbar sein, welche dem Diagramm (besser gesagt dem Benutzer) bekannt ist und möglichst eindeutig sein sollte. Hierfür bietet sich der Name der Shape an. Anstatt aber nun IDs in Form von Strings zu halten, sondern auf bestehende Objekte verweisen zu können, müssen die Instanzen der generierten Shapes usw. lokal in einer Map gesichert werden. Hierfür wurde eine Hilfsklasse *Cache* erstellt, die Container für alle relevanten Datentypen bereitstellt. So können fertig instantiierte Objekte zwischengespeichert und bei Bedarf über die entsprechende Map der Cache Instanz referenziert werden.

#### 4.5.1 Implicits

Hier wird Gebrauch von einer sehr mächtigen Eigenschaft von Scala gemacht, den *implicits*. Wenn man sich beispielsweise im Shapekonstruktor befindet und auf ein parentShape-element zugreifen will, bräuchte man normalerweise einen Aufruf wie den folgenden:

```
val someShape: Shape =
someCacheInstance.ContainerOfType[T].get("identifier").
  data
```

Oder weniger abstrakt:

```
val someShape: Shape= cache.shapeHierarchy(parent_ID).data
```

Selbst Aufrufe, die wie im obigen Beispiel recht kurz ausfallen, zerstören die Lesbarkeit ohnehin schon komplexer Codezeilen. Scala erlaubt es implicits zu definieren, welche unter anderem implizite Typumwandlungen nach angegebener Methode durchführen können. Wenn in einem sogenannten *package object* besagte implicits definiert werden, ist es sogar möglich das entsprechende package object in einer Quelldatei zu importieren und so die impliziten Definitionen

überall in den gewünschten Scope zu bringen. Im Falle der häufig benutzten Cache Klasse, sehen Zugriffe auf den Cache nun wie folgt aus:

```
val parentShape:Shape = parent_ID
```

`parent_ID` ist tatsächlich immernoch ein String, wird aber vom Compiler als unpassend erkannt. Daraufhin sucht er nach einer impliziten Methode, welche in der Lage ist, den String in den gewünschten Typ (hier Shape) umzuwandeln. Der Programmierer, kann seine Cache-Instanz also sogar ohne explizit auf sie zugreifen zu müssen - eben implizit - benutzen. So kann man sich auf die wesentlichen Aspekte der Problemstellung konzentrieren. In diesem Fall ist die manuelle Typisierung der `parentShape` Konstante allerdings zwingend, da der Compiler sonst eine valide Zuweisung eines Strings zu einer nicht typisierten Konstante feststellen würde und daraufhin den Typ String inferieren würde. Die entsprechende Definition des `implicit`s im package object `model.parser` sieht aus wie folgt:

```
implicit def IDtoShape(id:String)(implicit c:Cache):  
    Shape =  
c.shapeHierarchy(id).data
```

Die Umwandlung der mit `implicit` gekennzeichneten Funktion nimmt dem Programmierer nun den Zwischenschritt ab über die Cacheinstanz die richtige Collection ausfindig zu machen und den gewünschten String aufzulösen. Auffällig ist, dass gleich **zwei** `implicit`s vorkommen. Ersteres kennzeichnet die Methode für den Compiler, das zweite ist ein *implicit parameter*. Da der besagte Cache pro Parserinstanz erzeugt wird um statische Bindungen prinzipiell vermeiden zu können, kann in der Methodendefinition im Grunde auf keine Cache-Instanz zugegriffen werden. Das liegt daran, dass der Compiler ja nicht implizit weiß, in welcher Cache-Instanz er nach der ID suchen soll. Genau das, wird dem Compiler durch das zweite `implicit` mitgeteilt. Man sagt ihm quasi: *'Ich kennzeichne eine Cache-Instanz, nach der sollst du suchen'*. Diese zwei einfachen Tricks, verkürzen die Verschachtelung und steigern die Lesbarkeit des Codes. Zwar kann der Code für jemanden, der nicht weiß was `implicit`s sind, durchaus verwirrend aussehen, der Shape variable wird immerhin eindeutig einen String zuweisen. Doch der Code selbst, liest sich beinahe wie pseudo Code und der eigentliche Sinn erschließt sich dadurch besser. Entgegen der *dispatch* Kennzeichnung in Xtext, handelt es sich hierbei **nicht** um eine Art dynamische Erweiterung einer Klasse, sondern um eine implizite/automatisierte Typumwandlung. Anstatt also Methoden für einen Typ zu definieren, den dieser eigentlich garnicht besitzt, nur um ein Argument in der entsprechenden Parameterliste zu sparen, bleiben die Klassen hier konsistent. Im package object `parser` sind so für alle collections, die der Cache benutzt implizite Typumwandlungen bereitgestellt. So bleibt der Cache bis auf wenige Ausnahmen vollkommen unsichtbar für den Entwickler (wenn gewünscht). `Implicit`s sind zwar unwahrscheinlich mächtig, jedoch schwer wieder los zu werden (sollte dies je gewünscht werden). Deshalb muss vorsichtig überlegt werden, ob diese Vorgehensweise der späteren Objektstruktur hilft oder sie eher undurchsichtig macht.

# Kapitel 5

## Fazit

Abschließend zu der Arbeit werden Probleme benannt, die während des Prozesses aufgekommen sind. Außerdem wird zusammengefasst, was erreicht wurde, was eventuell versäumt wurde und welche Möglichkeiten sich in der Zukunft durch diese Kontribution für das MoDiGen Projekt ergeben.

### 5.1 Probleme

Die Arbeit war sehr zeitaufwendig, da die bestehende Objektstruktur, die durch die Xtext Grammatik beschrieben wurde auf Scala bezogen gelegentlich keinen oder wenig Sinn machte. Hin und wieder wurden auch Fehler in der definierten DSL entdeckt. Da das Umsetzen einer DSL in Scala ausgehend von einer Xtext Grammatik voraussetzt, dass entsprechende Grammatik verstanden wurde, dauerte es außerdem eine Weile, bis alle Unklarheiten bezüglich der Grammatik beseitigt wurden. Hin und wieder machten sich Fehler erst lange Zeit später bemerkbar, so mussten hin und wieder bestehende Klassen komplett neu überarbeitet werden. Auch Xtext Sprachkenntnisse waren nötig, da z.B. die Generatoren der Xtext Sprache sonst unklar erscheinen. Methodenaufrufe wie `createColorValue`, die wie in 4.1.2 beschrieben worden sind, sind ähnlich den *implicit*s in Scala für Neulinge absolut unverständlich. Auch bis die Xtext Quelldateien richtig interpretiert werden konnten, verging einige Zeit. Da das benötigte Metamodell zur Zeit der Entwicklung noch nicht einsatzbereit war, musste sich an einigen Stellen mit Mockups geholfen werden.

#### 5.1.1 Scalaspezifische Probleme

Scalaspezifische Probleme gab es tatsächlich nicht viele. Scala ist einfach zu erlernen, bietet komfortable Collections, sowie eine ausgezeichnete und einheitliche Collections API. Außerdem bietet Scala Features wie *case classes* und *Pattern Matching*, um nur ein paar zu nennen. Diese Features sind schnell verstanden und umsetzbar. Einzig und allein die Tatsache, dass hier ein Scala Neuling am Werk war, bremste die Arbeit erheblich aus. So wurde zunächst mit langen und sehr komplexen regulären Ausdrücken versucht, beliebige Shapedefinitionen zu parsen, was aber aufgrund der Rekursion der in sich geschach-

telten geometrischen Figuren unmöglich ist. Ein Scala Kenner, hätte sofort zu den *Parser Combinator* Klassen gegriffen, diese waren dem Autor zunächst aber noch nicht bekannt. Scala bietet viele Lösungen für viele Probleme, jedoch ist es insbesondere als Scala Neuling zunächst ein Labyrinth aus Möglichkeiten und oft wird erst später bemerkt, dass eine alternative Lösung besser passt. Ein tatsächlicher Nachteil von Scala gegenüber der Xtext Technik findet sich in der Implementierung der Generatorenklassen wieder. Xtext bietet für die Definition auszugebender Strings Makros an, welche es ermöglichen Kontrollstrukturen wie if-Bedingungen und Schleifen reibungslos in String Definitionen einzubringen. Scala verfügt zwar über einen ähnlich mächtiges Feature namens *String Interpolation*, jedoch hat dies einen entscheidenden Nachteil: Scope. Strings wie in Listing 5.1,

```

"""
Stencil.groups = {
    <<var i = 0>>
    <<val groupSet = getNodeToPaletteMapping(diagram).
        keySet()>>
    <<FOR group : groupSet>>
        <<getVarName(group)>> : {index: <<i=i+1>> ,
            label: '<<group>>' }<<IF group !=
                groupSet.last>> ,<<ENDIF>>
    <<ENDFOR>>
};
"""

```

Listing 5.1: String der in Xtext über Makros sauber formatiert erstellt werden kann

sind schon während sie entstehen so formatiert, wie ihr späterer Output. Das entsprechende Pendant in Scala ist in Listing 5.2 zu sehen.

```

var i = 0
var group = ""
val groupSet = getNodeToPaletteMapping(diagram).keySet
s"""
Stencil.groups = {
    ${
        for (groupName <- groupSet) yield {
            group = groupName
            getVarName(group)
        }
    }
    : {index: ${i = i + 1}, label: '$group' }${if (group !=
        groupSet.last) s", "}}
};

```



### Listing 5.2: String der in Scala über Interpolation modifiziert wird

Scala bietet es zwar an über *Escaping* `{}` Codeblöcke innerhalb von Strings auszuführen, jedoch eröffnen diese Codeblöcke auch einen eigenen Scope. Zählvariablen (beispielsweise) müssen in diesem Fall im überliegenden Scope definiert werden um alle separaten Code Blöcke im String beeinflussen zu können. Die entstandene Formatierung ist insgesamt nicht so gut wie die, die durch die Xtext Makros zu erreichen ist. Das ist tatsächlich nur ein kleines Ärgernis, aber der Vollständigkeit halber aufzuführen.

## 5.2 Schlussfolgerung/Ausblick

### 5.2.1 Was wurde erreicht

Die Objektstruktur wurde erfolgreich in Scala umgesetzt. Hierbei wurde darauf geachtet, dass alle Modellklassen immutabel sind. In Zukunft kann so auf funktionale Programmier Techniken zurückgegriffen und insbesondere paralleler Code effizient genutzt werden. Die DSLs zu Style, Shape, Connection und Diagram können erfolgreich eingelesen und übersetzt werden. Hierbei wurde darauf geachtet, dass eine erste Parserlogik (*Lexer*) die eingelesenen Strings in einheitliche Datenformate umwandelt, sodass spezialisiertere Parser ihren Input einheitlich behandeln können. So ist der Parser relativ leicht zu erweitern. Mehrfachvererbung bei Styles war mit Scala problemlos umsetzbar. Das iterieren über Elternelemente ist mittels generischer *Getter Funktionen* unwahrscheinlich einfach. Auch die (Mehrfach-) Vererbung bei Shapes funktioniert. Die Objektstruktur lässt sich also definitiv flexibler gestalten als in Xtext. Style Eigenschaften werden transitiv an unterliegende Elemente weitergegeben, wobei die hierdurch entstehenden Abhängigkeiten aus Anwendersicht statisch gebunden werden. Außerdem können abstrakte Shapes definiert werden, die ohne ein Diagramm erzeugt werden können um so allgemeine Eigenschaften einer logischen Gruppe zu bündeln. Geerbte Informationen werden prinzipiell zum Erzeugungszeitpunkt an Kindelemente weitergegeben. Somit ist die Anwendung besonders bei großen Baumstrukturen der Modellklassen nicht darauf angewiesen (eventuell auch mehrere male) Styleattribute über rekursive Methoden zu finden.

### 5.2.2 Was nicht erreicht wurde

Da zu Anfang entschieden wurde die Vergabe transitiver Style Eigenschaften schon bei Objekterzeugung weiterzugeben, musste später festgestellt werden, dass aufgrund der Instantiierungsreihenfolge Containerklassen notwendig sind, welche es erlauben die Erzeugung einer Modellklasse auf einen späteren Zeitpunkt zu verschieben. Dies erzeugt einen zusätzlichen Overhead an Speicherbedarf. Alle erzeugten Objekte von Modellklassen werden in einem Cache hinterlegt um zu ermöglichen sie über ihren Namen mittels eines Mappings zu referenzieren. Auch die Skizzen, also die Containerklassen, der Modellklassen müssen in diesem Cache hinterlegt werden, um fortlaufend bei Bedarf erzeugt werden

zu können. So können also die Containerklassen auch nicht nach Erzeugung einer zugehörigen Modellklasse gelöscht werden. Performanz für Speicherauslastung. Über die Anwendung *VisualVM* Kulkarni 2015, S.404-405 lässt sich ein *Heapdump* erstellen und analysieren. Beispielhaft wurden hierfür 10.000 Shapes eingelesen, um anschließend über VisualVM betrachtet werden zu können (siehe Abbildung 5.1). Ersichtlich wird, dass bei 10000 Shapes auch 10000 ShapeSket-

Class Name	Instances	Size
model.shapecontainer.shape.Shape	10,001 (0.5%)	1,360,136 (1.6%)
parser.ShapeSketch	10,000 (0.5%)	800,000 (1%)

Abbildung 5.1: Ausschnitt aus dem Heapdump, generiert durch JVisualVM

ches erzeugt werden. Das ist der angesprochene zusätzliche Overhead. Generell kann festgestellt werden, dass über die Lösung in Scala Funktionalität mit eingebracht werden konnte, die Xtext nicht bietet, jedoch bietet Xtext Vorteile die hier nicht verfolgt wurden, wie die Konsistenz zwischen DSL und Objektstruktur. Ändert sich in diesem Projekt die DSL oder die Objektstruktur, muss von Hand dafür gesorgt werden, dass Parser und Modellklassen konsistent bleiben. Das liegt daran, dass die DSL nur implizit über die Parser Regeln definiert ist.

### 5.2.3 Ausblick

Nächste Schritte für das Programm könnten sein, eine Scala Bibliothek zu erstellen, die eine DSL zur Beschreibung von Grammatiken anbietet. Das hört sich erst einmal verwirrend an. Derzeit wird über eine vordefinierte Grammatik eine (beispielsweise) Shape beschrieben und über einen Parser in ein bereits vordefiniertes Objekt umgewandelt. Da die Shape Klasse also wie gesagt schon vordefiniert ist besteht ein Konsistenzproblem zwischen DSL und Objektstruktur. Wird die Grammatik zur Beschreibung der Shape verändert, muss die Shape Klasse ebenfalls von Hand verändert werden um eine korrekte Abbildung zu garantieren. Würde eine DSL benutzt werden, die es erlaubt Grammatiken zu erstellen, könnten die bisher vordefinierten Klassen Style, Shape, Diagram über Generatoren aus der durch die DSL erstellte Grammatik erzeugt werden. So wäre eine konsistente Lösung geschaffen. Man müsste nicht sowohl Parserregeln als auch die Modellklassen verändern, sondern würde dies implizit tun, indem man die Definition der Modellklassen über die DSL verändert. Diese Vorgehensweise wird so in etwa auch von Xtext realisiert. Fraglich wäre also, ob dies so problemlos möglich ist, oder ob Features wie Mehrfachvererbung unter diesem Ansatz leiden könnten, immerhin ist dies so in Xtext nicht möglich. Des weiteren könnten Parsingvorgänge parallelisiert werden. Das würde die Performanz deutlich steigern und müsste auf Grund der immutablen Klassen auch einfach und ohne komplexe Synchronisationsmechanismen umsetzbar sein. Zusätzlich könnte überlegt werden, ob im Bezug auf die Vererbungsstrategie lediglich für Style Instanzen eine Ausnahme gemacht wird. Styles würden also nicht wie jetzt bei Objekterzeugung sondern erst, wenn Attribute abgefragt werden von den

Eltern ermittelt werden. Wie bereits gesagt muss hier unterschieden werden, ob schlussendlich mehrfaches rekursives Suchen und optimale Speicherauslastung oder eine nicht rekursive Problemstellung und ein damit einher gehender Overhead an Containerelementen bevorzugt werden. Vererbungsprinzipien sind nun gegeben und können falls gewünscht einfach auf Connection und Diagram übertragen werden. So könnten auch Connections und Diagrams von Elternelementen erben.

# Abbildungsverzeichnis

1.1	Programmierparadigmen im Verlauf der Zeit . . . . .	2
1.2	Klassendiagramm Diagram, Shape, Style (Sehr stark vereinfacht)	4
3.1	Klassenmodell von ClassHierarchy . . . . .	8
3.2	Sequenzdiagramm für den Ablauf eines Parsing Vorgangs . . . .	12
4.1	Vereinfachte Objektstruktur. Im Fokus Style und seine wichtig- sten Abhängigkeiten. . . . .	20
4.2	Vereinfachte Objektstruktur. Im Fokus Shape und seine wichtig- sten Abhängigkeiten. . . . .	22
4.3	Vereinfachte Objektstruktur. Im Fokus Diagram und seine wich- tigsten Abhängigkeiten. . . . .	23
4.4	Beispiel transitiver Styleinformation . . . . .	37
5.1	Ausschnitt aus dem Heapdump, generiert durch JVisualVM . . .	52

# Listings

2.1	Beispielhafte Variablendefinition mit Typinferierung und manueller Typangabe . . . . .	5
3.1	Beispielhafte Vereinfachung durch Apply Methode . . . . .	8
3.2	Beispielhafte Style Definition über die entsprechende Style DSL . . . . .	9
3.3	(Mehrfach)Vererbung bei einer Shape Definition . . . . .	9
3.4	vereinfachte Style Klassendefinition . . . . .	13
3.5	Style S1 . . . . .	13
3.6	Style S2 . . . . .	13
3.7	S1 und S2 Objekte (aus Gründen der Lesbarkeit aufgeführt in der <i>Java Script Object Notation</i> ) . . . . .	14
3.8	Definition einiger Styles, die vom jeweils nächsten Style erweitert werden . . . . .	15
4.1	stark vereinfachter Auszug aus der Xtext Grammatik, die Styles beschreibt . . . . .	18
4.2	Auszug aus der Style.xtext Grammatik, Definition der ColorOrGradient . . . . .	18
4.3	Auszug aus der Style.xtext Grammatik, Color kann als ColorOrGradient oder als ColorWithTransparency benutzt werden . . . . .	19
4.4	Auszug aus Code ColorConstants als case objects . . . . .	19
4.5	Auszug aus Code Trait ColorOrGradient . . . . .	19
4.6	Beispiel Ellipse die einer Stylerregel unterliegt und ein Rechteck beinhaltet . . . . .	21
4.7	Beispielnavigation durch komplexe Objektsruktur . . . . .	21
4.8	Beispielhafte Diagram Definition (vereinfacht) . . . . .	22
4.9	Auszug aus Code MclassMock über dass eigentliche MClass Elemente gemockt werden . . . . .	22
4.10	Auszug aus Code Viewbound. Implizite Umwandlung auf val name:String . . . . .	24
4.11	Auszug aus Code Apply Methoden, die die Benutzung von ClassHierarchy erleichtern . . . . .	24
4.12	Beispiel zur Instantiierung der abstrakten Klasse GradientAlignment . . . . .	26
4.13	Beispiel zur Instantiierung der abstrakten Klasse GradientAlignment über Vererbung . . . . .	26
4.14	Beispielhafte Shapedefinition . . . . .	27
4.15	Beispiel für eine rekursive Grammatik anhand der russischen <i>Babuschka</i> Puppen . . . . .	27

4.16	Einfaches Beispiel für einen Namens Parser . . . . .	28
4.17	Einfaches Beispiel zum Parsen einer rekursiven Grammatik . . .	28
4.18	Rekursive Methode zum Parsen geometrischer Figuren . . . . .	29
4.19	Auszug aus Code beispielhafte Regeln, die verallgemeinerte Attribut- Wert Tupel parsen . . . . .	30
4.20	Beispiel für eine Attributsdefinition mit expliziter Parameterbe- nennung . . . . .	30
4.21	Auszug aus dem Code zum Pattern Matching . . . . .	31
4.22	Auszug aus dem Code Methode ifValid . . . . .	31
4.23	Auszug aus Code sammeln der Elterninstanzen . . . . .	32
4.24	Auszug aus Code Zuweisung der Felder der Elternelemente . . .	32
4.25	Auszug aus Code Funktion mostRelevant . . . . .	33
4.26	Beispielhafter Aufruf von mostRelevant . . . . .	33
4.27	Auszug aus Code Hilfsfunktion relevant . . . . .	33
4.28	Beispielaufruf um die latest-Bound description der Elterninstan- zen zu ermitteln . . . . .	34
4.29	Erweiterung der geometrischen Figuren anhand z.B. der Tiefe der Elemente . . . . .	35
4.30	Einfügen seperater Bäume . . . . .	35
4.31	Beispiel zur Verdeutlichung transitiver Styleeigenschaften . . . .	36
4.32	Erzeugen eines neuen Styles über andere Styleinstanzen . . . . .	38
4.33	Auszug aus Code Erzeugen von Kindstyles . . . . .	38
4.34	Auszug aus Code Funktion generateChildStyle . . . . .	38
4.35	Auszug aus Code Definition der ShapeSketch container Klasse . .	39
4.36	Beispiel Aufrufe zum Provozieren einer möglicherweise ungewoll- ten Situation durch Vererbung . . . . .	41
4.37	Beispieldefinition für eine Shape . . . . .	43
4.38	Auszug aus Code Methode zum Parsen einer Shape(Sketch) . . .	43
4.39	Auszug aus Code Regeln zum Parsen der Felder . . . . .	44
4.40	Auszug aus Code Regel zum Parsen geometrischer Figuren . . .	44
4.41	Auszug aus Code Methode zum direkten Parsen von Shapes . . .	45
4.42	Auszug aus Code Iterieren durch Attribute mit Pattern Matching	45
4.43	Auszug aus Code Methode parseGeometricModels . . . . .	46
5.1	String der in Xtext über Makros sauber formatiert erstellt wer- den kann . . . . .	50
5.2	String der in Scala über Interpolation modifiziert wird . . . . .	50

# Literaturverzeichnis

- [1] A.V. Aho. *Compiler: Prinzipien, Techniken und Werkzeuge*. Pearson Studium Informatik. Pearson Education Deutschland, 2008. ISBN 9783827370976. URL <https://books.google.de/books?id=pTKAwL64NkoC>.
- [2] L. Bettini. *Implementing Domain-Specific Languages with Xtext and Xtend*. \*\*. Packt Publishing, Limited, 2013. ISBN 9781782160311. URL <https://books.google.de/books?id=ifaBCqGka2IC>.
- [3] Oliver Braun. *Scala Objektorientierte Programmierung*. Carl Hanser Verlag München, 2011.
- [4] F. Budinsky. *Eclipse Modeling Framework: A Developer's Guide*. The eclipse series. Addison-Wesley, 2004. ISBN 9780131425422. URL <https://books.google.de/books?id=ff-9ZYhvPwwC>.
- [5] Eclipse Modeling Framework (EMF). Eclipse modeling framework. <https://eclipse.org/modeling/emf/>. Accessed: 2016-02-06.
- [6] Friedrich Esser. *Scala für Umsteiger*. Oldenburg Verlag München, 2011.
- [7] Markus Gerhart and Prof Doc Marko Boger. Approach to define highly scalable metamodels based on json, .
- [8] Markus Gerhart and Prof Doc Marko Boger. Modigen-concept of a set of textual dsl to define graphical dsls based on a meta-model, .
- [9] M. Gunderloy and E. Tittel. *Developing XML Web Services and Server Components with Visual Basic .Net and the .NET Framework*. Exam cram 2. Que Certification, 2004. ISBN 9780789729002. URL <https://books.google.de/books?id=8Vh1yeokKZgC>.
- [10] R. Kulkarni. *Java EE Development with Eclipse*. Packt Publishing, 2015. ISBN 9781783550470. URL <https://books.google.de/books?id=xtVOCwAAQBAJ>.
- [11] M. Lange, W. Schubert, and I. Gringer. *Modellgetriebene Softwareentwicklung auf Basis von Xtext mit DSL*. 2013. URL <https://books.google.de/books?id=fdCHrgEACAAJ>.

- [12] R.C. Martin. *Clean Code: A Handbook of Agile Software Craftsmanship*. Pearson Education, 2008. ISBN 9780136083252. URL [https://books.google.de/books?id=\\\_i6bDeoCQzsC](https://books.google.de/books?id=\_i6bDeoCQzsC).
- [13] J. Trompeter and J.C.F. Beltran. *Modellgetriebene Softwareentwicklung: MDA und MDSD in der Praxis*. Entwickler.Press, 2007. ISBN 9783939084112. URL <https://books.google.de/books?id=iobrPQAACAAJ>.