

BACHELORTHESIS
Adrian Helberg

Template-basierte Synthese von Verzweigungsstrukturen mittels L-Systemen

FAKULTÄT TECHNIK UND INFORMATIK
Department Informatik

Faculty of Computer Science and Engineering
Department Computer Science

Adrian Helberg

Template-basierte Synthese von Verzweigungsstrukturen mittels L-Systemen

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung
im Studiengang *Bachelor of Science Angewandte Informatik*
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer: Prof. Dr. Philipp Jenke
Zweitgutachter: Prof. Dr. Michael Neitzke

Eingereicht am: 24. Februar 2021

Adrian Helberg

Thema der Arbeit

Template-basierte Synthese von Verzweigungsstrukturen mittels L-Systemen

Stichworte

Computergrafik, prozedurale Modellierung, 3D Generierung, L-Systeme, formale Grammatik, prozedurale generierung

Kurzzusammenfassung

Abstract: database amplification, gap between the model description and the control of its execution, Schwierigkeit den Expansionsprozess zielgerichtet zu kontrollieren, Prozedurale Regeln sind schwer zu verstehen und kontrollieren, Zugang, Klarheit schaffen, wie man an solche Fragestellungen herangeht, (Inverse) prozedurale Modellierung ist ein schwieriges Problem der Informatik → Quellen sichten und begründen warum das so schwierig ist

Adrian Helberg

Title of Thesis

Template-bases synthesis of branching structures using L-Systems

Keywords

Computer graphics, Procedural modeling, 3D generation, L-Systems, formal grammars, procedural generation

Abstract

The abstract belongs to here ...

Inhaltsverzeichnis

Abbildungsverzeichnis	vi
Tabellenverzeichnis	vii
Abkürzungen	viii
Symbolverzeichnis	ix
Listings	x
1 Einleitung	1
1.1 Problemstellung	2
1.2 Ziele	3
1.3 Methodik	3
1.4 Aufbau	5
2 Grundlagen	7
2.1 Grundbegriffe	7
2.2 Grundlegende Arbeiten	8
2.3 Verwandte Arbeiten	12
3 Konzepte	14
3.1 Probleme & Lösungsansätze	14
3.2 Workflows & Algorithmen	16
3.3 Softwarearchitektur	22
3.3.1 Grundkonzepte	22
3.3.2 Weitere Konzepte	28
3.3.3 Datenstrukturen	29
4 Implementierung	30
4.1 Projektstruktur	30

4.2	Technologien	32
4.3	Konzeptumsetzung	32
4.4	Entscheidungen	40
5	Evaluierung	41
5.1	Testumgebung	41
5.2	Beobachtungen & Ergebnisse	41
5.3	Diskussion und Bewertung	41
6	Ausblick	42
	Literaturverzeichnis	43
A	Anhang	47
	Glossar	49
	Selbstständigkeitserklärung	50

Abbildungsverzeichnis

1.1	Systemarchitektur	5
3.1	System und Systemumgebung	23
3.2	Interaktion zwischen System und Systemumgebung	23
3.3	Subsysteme mit fachlichen Abhängigkeiten	25
3.4	Subsystem GUI	26
3.5	Laufzeitsicht	27
3.6	Infrastruktur Windows-PC	27
4.1	Softwareprojekt Dateistruktur	31
4.2	Softwarepakete mit zugehörigen Funktionen	31
4.3	Startskript Generator.bat	33
4.4	Programm nach Ausführung des Start-Skripts	33
4.5	Erster Anker ist vorselektiert & gesetzte Parameter	34
4.6	Auswahl des ersten Templates	35
4.7	Der Verzweigungsstruktur hinzugefügte Template-Instanz	36
4.8	Vom Benutzer fertiggestellte Verzweigungsstruktur	36
4.9	Generierte Verzweigungsstrukturen	37
4.10	Verzweigungsstruktur & zugehörige Baumstruktur	39

Tabellenverzeichnis

Abkürzungen

HAW Hochschule für Angewandte Wissenschaften.

Symbolverzeichnis

Ω unit of electrical resistance.

Listings

3.1	Erstellen einer Verzweigungsstruktur	16
3.2	Inferieren eines L-Systems aus einer Baumstruktur	18
3.3	Erstellen eines kompakten L-Systems mit Gewichtung w_l	19
3.4	Kostenfunktion C_i mit Gewichtung w_l	19
3.5	Längenfunktion L für Grammatiken	20
3.6	Grammar Edit Distance	20
3.7	Kostenfunktion C_g mit Gewichtung w_0	20
3.8	Generalisieren eines L-Systems mit Gewichtung w_0	21

1 Einleitung

Hochschule für Angewandte Wissenschaften (HAW) Ω HAW Hamburg (Damit das glossar keine fehler schmeißt)

Effizientes Objekt design und -modellierung sind entscheidende Kernkompetenzen in verschiedenen Bereichen der digitalen Welt. Da die Erstellung geometrischer Objekte für Laien unintuitiv ist und ein großes Maß an Erfahrung und Expertise erfordert, ist dieses stetig wachsende Feld für Neueinsteiger nur sehr schwer zu erschließen. Die Forschung liefert hierzu einige Arbeiten zur prozeduralen Modellierung, um digitale Inhalte schneller und automatisiert zu erstellen. Gerade wenn es um die Darstellung natürlicher Umgebung geht, ist die Erstellung von ähnlichen Objekten, wie zum Beispiel verschiedene Bäume derselben Gattung eines Waldes, ein schwieriges Problem. Kleine Änderungen in prozeduralen Systemen können zu großen Veränderungen der Ergebnisse führen. Darum beschäftigt sich die inverse prozedurale Modellierung unter anderem mit dem Inferieren von Regeln und Mustern aus gegebenen Objekten, um diese nach bestimmten Regeln zu modellieren. Ein wichtiges Werkzeug hierbei ist die Verwendung formaler Grammatiken als fundamentale Datenstruktur der Informatik, um Strukturen zu beschreiben. Eine spezielle Untergruppe sind die L-Systeme, die häufig bei der Beschreibung von Verzweigungsstrukturen und Selbstähnlichkeit zum Einsatz kommen.

Diese Arbeit soll sich mit der Erstellung eines prozeduralen Systems zur Synthese von Ähnlichkeitsabbildern beschäftigen. Hierzu soll über eine Benutzerschnittstelle eine Struktur erzeugt werden, aus der ein parametrisiertes L-System inferiert werden kann, das dann zur Generierung von ähnlichen Strukturen genutzt werden kann.

1.1 Problemstellung

Mit der Digitalisierung der Welt steigt auch der Bedarf an digitalen Inhalten. Zu den größten Feldern gehören Gaming- und Unterhaltungsindustrie, Datenvisualisierung und interaktive Anwendungen. Um eine erhöhte Quantität dieser Inhalte liefern zu können, werden Methodiken und Algorithmen gesucht, die eine Erstellung vereinfachen. Während Methodiken zur Kodifizierung bestimmter Strukturen in den Bereich der prozeduralen Modellierung fällt, findet das Ableiten von Regeln Anwendung in der inversen prozeduralen Modellierung. Weiter steigt mit dem digitalen und naturwissenschaftlichen Fortschritt die Anwendung immer komplexerer Strukturen, die ein Herausarbeiten der schwer zu kontrollierenden, prozeduralen Regeln immer schwieriger machen.

Ein Beispiel hierzu aus der Gaming-Industrie ist die frühere Verwendung unorganisierter Modelle. Unorganisierte Modelle sind nur bedingt wiederverwendbar, da nur die vorliegende Modellierung verwendet werden kann. Es besteht eine hohe Speicherkomplexität bei geringer Laufzeitkomplexität. Für kleinste Veränderungen am Modell ist eine erneute Modellierung nötig, die wiederum Speicher benötigt, um sie persistent speichern zu können. Heutzutage werden die Objekte nach bestimmten Kriterien organisiert, um eine automatisierte Modellierung durch Algorithmen zu ermöglichen, um so aus einer Grundstruktur weitere Modelle zu erzeugen. Speicher- und Laufzeitkomplexität nähern sich an. Deshalb werden allgemeingültige, vielseitig anwendbare Algorithmen gesucht, die bestimmte natürliche Eigenschaften von Strukturen herausarbeiten („Reverse Engineering“), um diese für die (inverse) prozedurale Modellierung zur Verfügung zu stellen.

Diese Arbeit soll zeigen, wie sich durch die Erstellen eines Systems zur Generierung von ähnlichen Strukturen aus einer Basistruktur aktuelle Ansätze aus der Forschung in einem Programm umsetzen lassen.

1.2 Ziele

Die Erstellung eines Systems zur Synthese von ähnlichen Strukturen aus einer Basisstruktur soll zentrale Aufgabe dieser Arbeit sein. Aus der Fragestellung leiten sich folgende Teilziele ab:

- Die Erstellung eines Programms zur Umsetzung des erstellten Systems
- Anwenden von Algorithmen und Ansätzen der aktuellen Forschung
- Testen von Metriken und deren Auswirkung auf das Ergebnis
- Erstellung eigener Methodiken und Algorithmen zur Effizienten Lösung der Problemstellung
- Schaffen eines Teilsystems zum Extrahieren von Eigenschaften und Regeln einer Basisstruktur

Die Erstellung und Integration eines neuronalen Netzes, das laut aktueller Forschung gute Ergebnisse beim Lernen von Regeln aus einer Eingabestruktur liefert, kann in dieser Arbeiten aus Praktikabilitätsgründen nicht behandelt werden.

1.3 Methodik

Der Benutzer des Systems legt atomare Strukturen in Form von Zeichenketten an, die vom Programm eingelesen und als Templates zur Verfügung gestellt werden. Die Templates sind beliebig und können eine einfache Linie oder eine komplexe Verzweigung darstellen. Werden diese Strukturen auf der graphischen Oberfläche platziert und mit diversen Transformationen verändert, spricht man von Instanzen oder Template-Instanzen.

Strukturieren

Der Benutzer nutzt die grafische Benutzerschnittstelle, um aus einzelnen Templates eine zusammenhängende Basisstruktur zu erstellen. Neben der Position der Instanzen werden Transformationsparameter, wie Rotation oder Skalierung, angepasst.

Visualisieren

Der aktuelle Stand der Strukturierung ist jederzeit sichtbar. Liniensegmente und Bindungselemente werden in einem graphischen Element visualisiert und für eine Interaktion zur Verfügung gestellt.

Datenaufbereitung

Das Ergebnis der Strukturierung wird in einer Baumstruktur organisiert, in der jeder Knoten einer bestimmten Template-Instanz entspricht und die eingehenden Kanten die geometrischen Transformationen relativ zum Elternknoten beschreibt.

Inferieren

Aus der Baumstruktur kann eine formale Grammatik in Form eines L-Systems abgeleitet werden. Diese Grammatik beschreibt lediglich die erstellte Basisstruktur und beinhaltet keine Transformationsparameter, da hier nur auf die Topologie des Baumes und nicht auf geometrische Unterschiede der Instanzen eingegangen wird.

Komprimieren

Um sich wiederholende Produktionsregeln zu vermeiden und so sowohl das Alphabet, als auch die Produktionsregelmenge zu kompromieren, wird die Baumstrukturen nach identischen, maximalen Unterbäumen durchsucht und durch zusammengefasste Instanzen vereinfacht. Hierbei gilt die Baumstruktur selbst nicht als Unterbaum.

Generalisieren

Ähnliche Produktionsregeln des L-Systems werden mithilfe einer Kostenfunktion zusammengefasst, um dies mit nicht-deterministischen Regeln und Rekursion zu generalisieren.

Randomisieren

Jedes Symbol der Grammatik nimmt eine Liste an Parametern entgegen, die nach bestimmten Kriterien pseudo-randomisiert werden, um Variationen von Template-Instanzen zu erstellen. Das Ausführen des L-Systems kann nun Ähnlichkeitsstrukturen für die Basisstruktur erzeugen.

1.4 Aufbau

Die Methodik zum Umsetzen des beschriebenen Systems wird wie folgt umgesetzt.

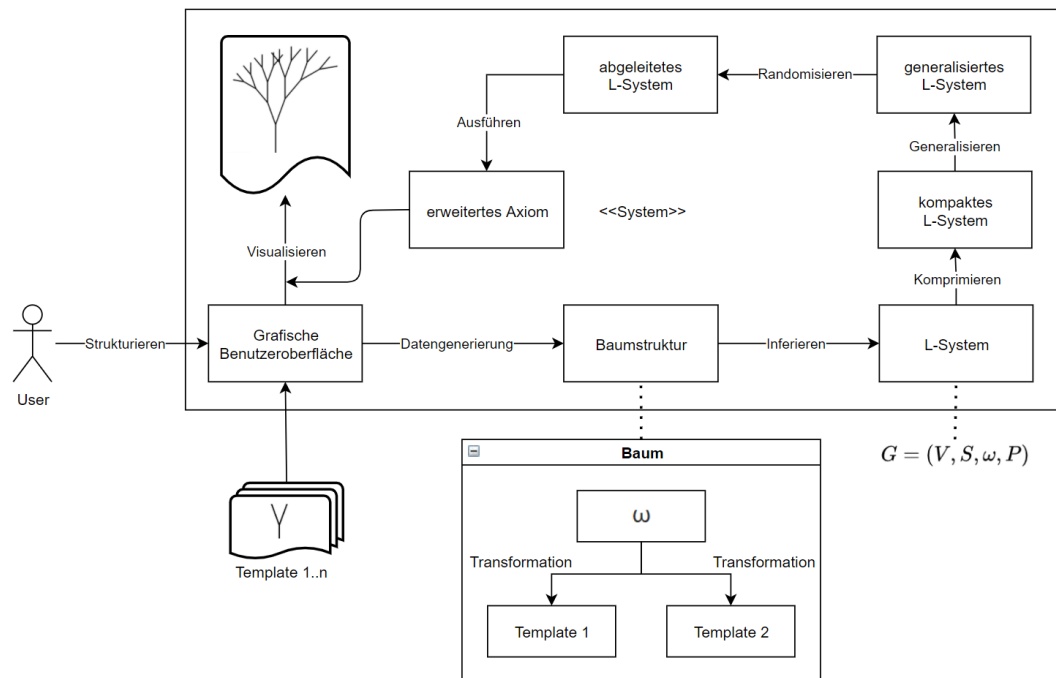


Abbildung 1.1: Architektur des Systems mit einigen Datenstrukturen

Die Darstellung zeigt eine grobe Übersicht der Anwendungsfälle innerhalb des Systems. Der Benutzer strukturiert vom System importierte Templates zu einer Basisstruktur mittels grafischer Bedienelemente. Die Template-Instanzen werden hierbei in einer Baumtopologie organisiert. Anschließend wird ein L-System aus der Baumstruktur generiert und komprimiert. Bevor randomisiert, geometrische Parameter auf das kompakte L-System

angewendet werden, wird dieses generalisiert. Zuletzt wird das Axiom des abgeleiteten L-Systems durch die Ausführung erweitert und sichtbar gemacht.

2 Grundlagen

Die Modellierung mithilfe von Grafiksoftware ist eine vergleichbar händische, langwierige Erstellung von Objekten. Hierbei hat der Designer (Modellierer) die volle Kontrolle über die Strukturen des Objektes.

Bei der prozeduralen Modellierung werden spezifische Strukturen eines zu erstellenden physikalischen Objektes generalisiert und meist über eine Grammatik und globale Parameter abgebildet. Während bei der klassischen Modellierung die menschliche Intuition und bei der prozeduralen Modellierung eine parametrisierte Grammatik vorausgesetzt wird, arbeitet die inverse prozedurale Modellierung mit bestehenden Modellen und extrahiert („lernt“) die Strukturen des Objektes, die automatisch in eine formale Grammatik überführt werden können. Die Generierung von prozeduralen Modellen ist ein wichtiges, offenes Problem [4].

2.1 Grundbegriffe

Modellierung

Um einen physikalischen Körper in ein digitales Objekt zu überführen, wird mithilfe von Abstraktion (Modellierung) ein mathematisches Modell erstellt, das diesen Körper formal beschreibt. 3D Grafiksoftware, wie bspw. Blender [30], wird genutzt um geometrische Körper zu modellieren, texturieren und zu animieren.

Prozedurale Modellierung

It encompasses a wide variety of generative techniques that can (semi-)automatically produce a specific type of content based on a set of input parameters [24]

Prozedurale Modellierung beschreibt generative Techniken, die (semi-)automatisch spezifische, digitale Inhalte anhand von deskriptiven Parametern erzeugen (*Übersetzt durch den Autor*). Smelik u. a. beschreibt einen Prozess, welcher durch das Nutzen globaler Parameter und descriptiven Regeln Modelle erzeugt [24].

Inverse prozedurale Modellierung

Aliaga u. a. spricht bei der inversen prozeduralen Modellierung von dem Finden einer prozeduralen Repräsentation von Strukturen bestehender Modelle [2]. Die Methodik aus Strukturen bestimmte Regeln und Parameter abzuleiten ist der Hauptgegenstand dieses Feldes der Computergrafik und aktueller Gegenstand der Forschung.

2.2 Grundlegende Arbeiten

Smelik u. a. untersucht prozedurale Methoden, um diverse Strukturen, wie Vegetation, Straßen u.v.m zu erzeugen. Es wird ein Überblick aktueller, vielversprechender Studien gegeben, deren Anwendung sowohl in technischen Bereichen, als auch in nicht-technischen, kreativen Bereichen, diskutiert wird. [24] gilt als grundlegender Einstiegspunkt in die Bereiche der prozeduralen Modellierung. Einen aktuellen Stand der Forschung der inversen prozeduralen Modellierung liefert [2]. Aliaga u. a. zeigt, dass IPM-Ansätze in entsprechende Kernprobleme der Informatik aufgeteilt und meist getrennt voneinander durch verschiedene Methodiken und Algorithmen bearbeitet werden. Weiter wird ein Einblick in die Kategorisierung und Bewertung einiger Ergebnisse von IPM-Systemem gegeben. Die Buchrezension [6] geht auf die Arbeit [12] ein und weist darauf hin, dass De la Higuera einen wesentlichen Unterschied zwischen der Induktion einer Grammatik, also das Finden einer Grammatik, welche ein Datum am genauesten beschreibt, und der Grammatikinferrierung, also das Finden einer Zielgrammatik, welche eine bestimmte Zeichenfolge abdeckt, sieht. De la Higuera ordnet die inverse Generierung von L-Systemen zum Problem der Grammatikinferrenz, welches er als gut erforschtes Gebiet beschreibt. Verzweigungsstrukturen im Kontext der inversen prozeduralen Modellierung tauchen in wissenschaftlichen Studien wenig auf. [11] adressiert diese in seiner Arbeit [11] und liefert einige Ansätze hierzu.

Im Folgenden wird eine Übersicht zu einigen Arbeiten zur Modellierung von bestimmten Strukturen gegeben.

Subjekt	Arbeit
Bäume und Landschaften	[7]
Fassaden	[1]
Gebäude	[18]
Städte	[21]
Möbel	[17]
Inneneinrichtung	[32]

L-Systeme

Lindenmayer führt eine mathematische Beschreibung zum Wachstum fadenförmiger Organismen ein. Sie zeigt, wie sich der Status von Zellen infolge ein oder mehrerer Einflüsse verhält [14]. Weiter führt er Ersetzungssysteme ein, die atomare Teile mithilfe von Produktionsregeln ersetzen. Diese L-Systeme nutzt er zur formalen Beschreibung von Zellteilung. Später werden Symbole zur formalen Beschreibung von Verzweigungen, die von Filamenten abgehen, eingeführt [23]. Die bekanntesten L-Systeme sind zeichenkettenbasiert und werden von *Noam Chomsky* eingeführt [5]. Sie ersetzen parallel Symbole eines Wortes, die von einer Grammatik über eine Sprache akzeptiert werden. L-Systeme können unter anderem parametrisiert oder nicht-parametrisiert und kontextfrei oder kontextsensitiv sein.

L-Systeme sind Grammatiken mit folgender Form:

$$\mathcal{L} = \langle M, \omega, R \rangle, \text{ mit}$$

- M als Alphabet, das alle Symbole enthält, die in der Grammatik vorkommen,
- ω als Axiom oder „Startwort“ und
- R als Menge aller Produktionsregeln, die für \mathcal{L} gelten

Das Alphabet eines parametrisierten Systems enthält Module (Symbole mit Parametern) anstatt Symbole:

$$M = \{A(P), B(P), \dots\} \text{ mit}$$

- $P = p_1, p_2, \dots$ als Modulparameter

Zeichen des Alphabets, die Ziel einer Produktionsregel sind, heißen Variablen. Alle anderen Zeichen aus M sind Konstanten. Das Axiom ω ist eine nicht-leere Sequenz an Modulen aus M^+ mit

- M^+ als Menge aller möglichen Zeichenketten aus Modulen aus M

Produktionsregeln sind geordnete Paare aus Wörtern über dem Alphabet, die bestimmte Ersetzungsregeln umsetzen. Hierbei werden Symbole aus einem Wort, die einer rechten Seite (*engl. right hand side (RHS)*) einer Produktionsregel entsprechen, durch die linke Seite des Paares (*engl. left hand side (LHS)*) ersetzt. Sie sind folgendermaßen aufgebaut:

$$A(P) \rightarrow x, x \in M^*, \text{ mit}$$

- M^* als die Menge aller möglichen Zeichenketten von M inklusive der leeren Zeichenkette ε

Ist die RHS jeder Produktionsregel ein einzelnes Symbol und gibt es zu jeder Variablen eine Regel, spricht man von einem kontextfreien, andernfalls von einem kontextsensitiven L-System.

Logo-Turtle-Algorithmus

Der Logo-Turtle-Algorithmus [22] setzt ein Vorgehen zur graphischen Beschreibung von L-Systemen, bei dem jeder Buchstabe in einem Wort einer bestimmten Zeichenoperation zugewiesen wird, um. So kann aus einem L-System ein grafisches Muster generiert werden, das mit einer Abfolge von Zeichenbefehlen an eine „Schildkröte“ gezeichnet wird. Das Triplett (x, y, θ) definiert den Status (State) der Schildkröte. Dieser setzt sich aus der aktuellen Position $\begin{pmatrix} x \\ y \end{pmatrix}$ und dem aktuellen Rotationswinkel θ , der die Blickrichtung bestimmt, zusammen.

Der Algorithmus kann als Komprimierung eines geometrischen Musters gesehen werden. Folgende Symbole mit zugehörigen Steuerungsbefehlen und Statusveränderung sind definiert:

Symbol	Steuerung	Statusveränderung
$F(d)$	Gehe vom derzeitigen Punkt p_1 d Einheiten in die Blickrichtung zu dem Punkt p_2 . Zeichne ein Liniensegment zwischen p_1 und p_2	ja
$+(\alpha)$	Setze neuen Rotationswinkel $\theta = \theta + \alpha$	ja
$-(\alpha)$	Setze neuen Rotationswinkel $\theta = \theta - \alpha$	ja
[Lege den aktuellen State auf einen Stack	nein
]	Hole den State vom Stack und überschreibe den aktuellen mit diesem	nein

Alles zwischen den Symbolen [und] wird als Verzweigung interpretiert.

Bsp. $FF[FF]F$ mit Verzweigung $[FF]$

Extreme Programming

Eine Fallstudie der Universität Karlsruhe [19] untersucht den Einsatz der Softwaretechnik Extreme Programming (XP) im Kontext der Erstellung von Abschlussarbeiten im Universitätsumfeld. Hierzu werden folgende Schlüsselpraktiken untersucht:

- XP als Softwaretechnik zur schrittweisen Annäherung an die Anforderungen eines Systems
- Änderung der Anforderungen an das Systems
- Funktionalitäten (Features) werden als Tätigkeiten des Benutzers (User Stories) definiert
- Zuerst werden Komponententests (Modultests) geschrieben und anschließend die Features (Test-driven Design)

- Keine separaten Testing-Phasen
- Keine formalen Reviews oder Inspektionen
- Regelmäßige Integration von Änderungen
- Gemeinsame Implementierung (Pair Programming) in Zweiergruppen

Aus der Fallstudie geht hervor, dass Extreme Programming einige Vorteile bei der Bearbeitung eines Softwareprojektes einer Bachelorarbeit bietet. Zum Einen können sich Anforderungen an das zu erstellende System durch parallele Literaturrecherche ändern, zum Anderen können Arbeitspakete durch Releases abgebildet werden.

2.3 Verwandte Arbeiten

Guo u. a. führt ein Modell zum Lernen von L-Systemen von Verzweigungsstrukturen mithilfe maschinellen Lernens (Deep Learning) anhand beliebiger Grafiken ein [11]. Hierzu werden atomare Strukturen mit einem neuronalen Netz erkannt, eine hierarchische Topologie (Baumstruktur) aufgebaut, aus der ein L-System inferiert und mit einem Greedy Algorithmus optimiert wird. Ausgabe des Systems ist ein generalisiertes L-System, aus dem ähnliche Strukturen, wie die der Inputgrafik, erstellt werden können.

Das Lernen von Design-Patterns mithilfe von Bayes-Grammatiken ist Gegenstand der Arbeit von Talton u. a.. [28] führt ein System zur Generierung geometrischer Modelle und Websites ein, das eine organisierte Struktur von bezeichneten Teilmodellen entgegennimmt und über einen Prozess der MCMC-Optimierung eine Bayes-optimale Grammatik erstellt, um neue Modelle zu generieren.

Auch [25] etabliert einen MCMC-Ansatz zum Finden prozeduraler Repräsentationen für biologische Bäume. Hier wird zunächst über eine Laplace-Glättung ein Grundgerüst gefunden, das dann in einer Baumtopologie organisiert wird.

Das von Martinovic und Van Gool eingeführte System nutzt ähnlich organisierte Eingabestrukturen in Form von Gebäudefassaden, um durch bayesische Grammatikinduktion eine kontextfreie Grammatik zu induzieren [15]. Das Erstellen kontextfreier Grammatiken mithilfe statistischer Methoden zur Verteilung von zweidimensionalen Clustern wird in der Arbeit von Stava u. a. gegeben

urbanen Strukturen [20] Polynomiale Algorithmen zum Inferrieren von L-Systemen [16]
Framework zum Inferrieren von L-Systemen aus Vektordaten [26] Generalisieren einen
Regelsets aus einem Inputset mit Markov Chain Monte Carlo [28, 27] Lernen von Regeln
zum Layout für Gebäude [15] Inverse prozedurale Modellierung von Fassaden [31]

3 Konzepte

Mit der Erstellung eines Programms zur Synthetisierung von Ähnlichkeitsabbildungen einer vom Benutzer erstellten Verzweigungsstruktur soll die Praktikabilität aktueller Forschungsansätze untersucht werden.

Folgende Kernkonzepte werden erläutert und umgesetzt:

- Visualisierung und prozessorientiertes Erstellen von Basisstrukturen,
- Organisation in einer prozessoptimierten, baumähnlichen Topologie,
- L-System-Repräsentationen,
- Algorithmen zur Inferierung, Komprimierung, Generalisierung und
- Verarbeitung von Transformationsparametern

3.1 Probleme & Lösungsansätze

Visualisierung

Um eine geführte Erstellung der Basisstruktur zu ermöglichen, muss diese während der Erstellung sichtbar gemacht werden. Hierzu werden die Templates in Form von Zeichenketten angelegt und mittels Turtle-Grafik visualisiert. Eine Turtle-Grafik beschreibt die Interpretation einer Zeichenkette als Bild durch Ausführen eines Logo-Turtle-Algorithmus. Weiter wird auch zur Evaluation von Ergebnissen eine Visualisierung benötigt. Da die Verzweigungsstrukturen in L-System-Repräsentation vorliegen, wird hierzu eine Interpretationsfunktion benötigt, die diese Ersetzungssysteme in Bildform darstellen. Ein L-System wird durch Ausführung in eine erweiterte Zeichenkette überführt und als Turtle-Grafik beschrieben [22].

Basisstruktur

Der Benutzer nutzt grafische Bendienelemente, um eingeleseene Templates auszuwählen, Transformationsparameter anzupassen und um anschließend die Instanzen der Basisstruktur hinzuzufügen. Im Folgenden wird diese Basisstruktur u.a. Grundstruktur und Eingabestruktur genannt.

Baumstruktur

Um Grundstrukturen mittels verschiedener Algorithmen untersuchen zu können, werden die einzelnen Template-Instanzen in einer baumähnlichen Struktur organisiert. Transformationsparameter einer Instanz beschreiben die räumlichen Veränderung gegenüber des zugrundeliegenden Templates und haben daher keine Aussagekraft in Bezug auf die Strukturtopologie der Basisstruktur. Diese Arbeit fokussiert sich auf topologische Eigenschaften von Verzweigungsstrukturen (z.B. Rekursionen). Darum bilden die einzelnen Template-Instanzen die Knoten der Baumtopologie, während die Kanten die räumlichen Transformationen darstellen. So wird eine datenstrukturelle Trennung zwischen Topologie und räumlichen Transformationen geschaffen. Diese baumähnliche Struktur ist inspiriert durch [11], Kapitel 4.2 *Grammar inference*

Inferieren

Das Smallest Grammar Problem, also das Finden der kleinsten, kontextfreien Grammatik, welche eine bestimmte Zeichenkette generiert, ist ein offenes Problem der Informatik mit einem Annäherungsverhältnis von weniger als $\frac{8569}{8568}$ (NP-hard). Primär wird in der Forschung nach Algorithmen gesucht, die ein akzeptables Ergebnis liefern. In dieser Arbeit wird ein Algorithmus präsentiert, der die Knoten der Baumstruktur in einzelne Symbole umwandelt, mit Produktionregeln verknüpft und diese dem resultierenden L-System hinzufügt. Dieses L-System repräsentiert lediglich die Eingabestruktur.

Komprimieren

Um ein kompaktes, gewichtetes L-System zu erzeugen, werden sich wiederholende Unterräume gesucht und ersetzt. Eine Gewichtung wird angewendet, um das zu erzeugende

L-System mit kleiner Regelmenge oder mit großer Regelmenge auszustatten. Eine Kostenfunktion stellt hierbei die Anzahl Symbole aller RHS der Produktionsregeln mit der Menge an Anwendungen der LHS gegenüber.

Generalisieren

Da das kompakte L-System eine Repräsentation der vom Benutzer erzeugten Verzweigungsstruktur darstellt, werden ähnliche Regeln miteinander verbunden (Merge) und mit einer Wahrscheinlichkeit versehen, um nicht-deterministische Regeln hinzuzufügen. Eine weitere Kostenfunktion bewertet den Merge zweier Produktionsregeln und wendet eine Gewichtung über die Länge der Grammatik zur Änderungsdistanz der alten (ohne Merge) zur neuen Grammatik um.

Transformationen

Die vom Benutzer vergeben Werte der Transformationsparameter werden während der Erstellung der Eingabestruktur in einer Häufigkeitsverteilung organisiert und bei Ausführung des generalisierten L-Systems angewendet. So werden Transformationen nach ihrer statistischen Häufigkeit angewendet.

3.2 Workflows & Algorithmen

Verzweigungsstruktur erstellen

Um als Benutzer des Systems eine Verzweigungsstruktur zu erstellen, wird folgender Arbeitsablauf umgesetzt:

Algorithmus 31: Erstellen einer Verzweigungsstruktur

- 1 Erster Anker ist vorselektiert
 - 2 Wiederhole, bis Struktur fertiggestellt ist:
 - 3 Selektiere ein Template aus der Liste
 - 4 Setzte Parameter
 - 5 Bestätige Auswahl und Parameter
 - 6 Zeichne ausgewähltes Template mit Parametern
 - 7 Wähle nächsten Anker aus
-
-

L-System inferrieren

Aus der Verzweigungsstruktur kann nun ein L-System erzeugt werden. Hierzu wird ein neuer Algorithmus präsentiert:

Zunächst werden L-System-Komponenten initialisiert (Zeile 2-6) für $\mathcal{L} = \langle M, \omega, R \rangle$:

- Das Alphabet wird mit den Symbolen F und S initialisiert, da F als Repräsentation einer grundlegenden Zeichenoperation und S als Axiom in jeder Anwendung des Algorithmus vorkommt
- Die erste Produktionsregel α umfasst die Abbildung des Axioms auf einen neues Symbol, das nicht im Alphabet vorkommt. Das Alphabet wird stets um ein Symbol lexikografischer Ordnung ergänzt
 - Bsp.: $\{A, B, C\}$ wird ein neues, unbekanntes Symbol hinzugefügt $\rightarrow \{A, B, C, D\}$
- Die Variable β hält zu untersuchende Knoten der Baumtopologie, welche nach Breitensuche iteriert werden. Die erste Iteration startet bei Wurzelknoten S.
- Als letzten Schritt der Initialisierung wird dem Alphabet ein neues Symbol hinzugefügt, das durch die Variable γ gehalten wird.

Die Schleife des Algorithmus beschäftigt sich mit der Iterierung des Baumes und dem Erstellen neuer Symbole und Produktionsregeln für das resultierende L-Systems (Zeile 8-17):

- Die in den Knoten des Baumes gehaltenen Template-Instanzen entsprechen einer Zeichenkette, die durch eine Turtle-Grafik interpretiert, einem vom Benutzer transformierten Template entspricht. Diese Zeichenkette wird in δ gespeichert
- In Zeile 9-12 wird die genannte Zeichenkette auf Verzweigungsvariablen ($A - Z$; F ausgeschlossen) untersucht. Diese werden durch ein neues Symbol, das dem Alphabet hinzugefügt wird, ersetzt. Anschließend wird eine Produktionsregel, die auf die veränderte Zeichenkette abbildet, der Produktionsregelmenge hinzugefügt.
- Zeile 13 prüft, ob es Symbol im Alphabet gibt, das nicht als Ziel einer Produktionsregel in der Produktionsregelmenge definiert ist. Ist dies der Fall, wird dieses als Ziel der nächsten Produktion gesetzt. Andernfalls schließt der Algorithmus ab

- Schleifen-Attribut ist hier das Setzen des nächsten Knotens am Ende der Schleife (Zeile 17)

Algorithmus 32: Inferieren eines L-Systems aus einer Baumstruktur

```

1   Initialisierung :
2    $M = \{F, S\}$ 
3    $\omega = S$ 
4    $R \leftarrow \{\alpha: S \rightarrow A\}$ 
5    $\beta = \text{nächster Knoten}$ 
6    $M \leftarrow \gamma \in \{A, B, \dots, Z\}, \text{ mit } \gamma \notin M$ 
7   Schleife :
8    $\delta = \text{Wort von } \beta$ 
9    $\forall \{A, B, \dots, Z\} \setminus F \in \delta :$ 
10  Ersetze mit  $\zeta \in \{A, B, \dots, Z\}, \text{ mit } \zeta \notin M$ 
11   $M \leftarrow \zeta$ 
12   $R \leftarrow \{\gamma \rightarrow \delta\}$ 
13  Wenn es ein Symbol  $\eta$  in  $M \setminus \{F, S\}$  gibt mit  $\{\eta \rightarrow bel.\} \notin R :$ 
14   $\gamma = \eta$ 
15  Sonst :
16  Breche Schleife ab
17   $\beta = \text{nächster Knoten}$ 

```

L-System komprimieren

Guo u. a. führt einen Algorithmus zum Inferieren einer Grammatik aus einer Baumstruktur ein [11]. Zum Einen wird ein L-System aufgebaut, zum Andern die Baumtopologie durch Finden maximaler Subbäume reduziert. Die Reduktion wird im folgenden Algorithmus adaptiert:

Initialisierung (Zeile 2-5):

- Das L-System, welches der Eingabe des Algorithmus entspricht, wird ausgeführt und die resultierende Zeichenkette wird in \mathcal{L}^+ gespeichert
- Ein Gewichtungssparameter w_l wird eingeführt, der eine Kostenfunktion (Zeile 11) nach Anzahl an Symbolen der RHS von Produktionsregeln und Anzahl deren Anwendung gewichten soll.
- Anschließend wird ein maximaler Subbaum durch geschachtelte Iteration des Baumes T gesucht und als T' gesetzt

- Es werden ausschließlich maximale Unterbäume behandelt, die mindestens zwei mal im Baum vorkommen

Die Schleife (Zeile 7-13) stellt die Reduzierung dar:

- Zunächst werden alle Vorkommen des maximalen Unterbaumes durch ein neues Symbol ersetzt
- Aus diesem Subbaum wird ein L-System inferriert, das wiederum in eine erweiterte Zeichenkette ausgeführt wird
- Die Zeichenkette wird als LHS einer neuen Produktionsregel gesetzt, die auf das neue Symbol abzielt
- Das alte L-System kann nun mit dem veränderten L-System mittels Kostenfunktion verglichen werden (Zeile 9): Liegen die Kosten des veränderten L-Systems unter den Kosten des Alten, wird die Reduktion beendet. Andernfalls gilt der Subbaum nun als Eingabebaum und das veränderte Ersetzungssystem als Eingabe-L-System

Algorithmus 33: Erstellen eines kompakten L-Systems mit Gewichtung w_l

```

1 Initialisierung :
2  $\mathcal{L}^+ \leftarrow L_s$ 
3  $\mathcal{L} = \emptyset$ 
4 Setze Gewichtungssparameter  $w_l \in [0, 1]$ 
5 Finde maximalen Unterbaum  $T'$  aus  $T$  mit Wiederholungen  $n > 1$ 
6 Schleife (Reduzierung):
7 Ersetze alle Vorkommen von  $T'$  mit dem selben Symbol  $\gamma \in \{A, B, \dots, Z\}$ 
8  $R \leftarrow \{\gamma \rightarrow L_s\}$  mit  $L_s$  aus  $T'$ ,  $R$  aus  $\mathcal{L}$ 
9 Wenn  $C_i(\mathcal{L}) \geq C_i(\mathcal{L}^+)$ 
10 Breche Schleife ab
11  $T \leftarrow T'$ 
12  $\mathcal{L}^+ \leftarrow \mathcal{L}$ 
13 Finde maximalen Unterbaum  $T'$  aus  $T$  mit Wiederholungen  $n > 1$ 

```

Algorithmus 34: Kostenfunktion C_i mit Gewichtung w_l

```

1  $C_i(\mathcal{L}) = \sum_{A(P) \rightarrow M^* \in \mathcal{L}} w_l * |M^*| + (1 - w_l) * N(A(P) \rightarrow M^*)$ 

```

mit $N(\cdot)$ als Zählfunktion für die Anzahl Wiederholungen einer *LHS* einer Regel in einem ausgeführten L-System.

L-System generalisieren

Da das kompakte L-System eine Repräsentation der vom Benutzer erzeugten Verzweigungsstruktur darstellt, werden nun ähnliche Regeln miteinander verbunden und mit einer Wahrscheinlichkeit versehen, um nicht-deterministische Regeln hinzuzufügen. Sowohl die Längenfunktionen und Kostenfunktionen, als auch der Grundalgorithmus sind aus [11] entnommen und bauen sich wie folgt auf:

- Die Längenfunktion L , die auf eine Grammatik angewendet wird, summiert die Anzahl Symbole des Alphabets mit der Anzahl an RHS der Produktionsregeln und misst somit die Gesamtheit aller Symbole, die das L-System abbilden soll (Länge der Grammatik)
- Der Abstand zweier Zeichenketten kann mit der *String Edit Distance* ermittelt werden. Diese wird über die Funktion D_s abgebildet. Hierbei wird die Anzahl an Operationen summiert, die für die Überführung einer Zeichenkette in eine Andere nötig sind (Zeichenkettenaustausch, -einschub und -löschung)
- Mit D_s kann nun auch der Abstand zweier Grammatiken zueinander bestimmt werden. Diese Funktion wird mit D_g abgebildet
- Die Kostenfunktion C_g nutzt die Länge und Distanz von Grammatiken, um Kosten einer Überführung einer Grammatik in eine andere messen zu können

Algorithmus 35: Längenfunktion L für Grammatiken

$$1 \quad L(\mathcal{L}) = |M| + \sum_{A(P) \rightarrow M^* \in \mathcal{L}} |M^*|$$

Algorithmus 36: Grammar Edit Distance

$$1 \quad D_g(\mathcal{L}^+, \mathcal{L}^*) = \sum_{(A(P) \rightarrow M_A^*, B(P) \rightarrow M_B^*) \in M(\mathcal{L}^+ \rightarrow \mathcal{L}^*)} D_s(M_A^*, M_B^*)$$

Algorithmus 37: Kostenfunktion C_g mit Gewichtung w_0

$$1 \quad C_g(\mathcal{L}^*, \mathcal{L}^+) = w_0 * (L(\mathcal{L}^*) - L(\mathcal{L}^+)) + (1 - w_0) + D_g(\mathcal{L}^+, \mathcal{L}^*)$$

Algorithmus zum Generalisieren eines L-Systems:

Initialisierung (Zeile 2-4):

- Das zu untersuchendes Tuple bestehend aus zwei Produktionsregeln (Regelpaar) und wird in der Variable p^* gehalten
- L-Systeme, die sich infolge eines Merges geändert haben, werden in L^+ und Eingabe-L-systeme in L^* gespeichert
- Die Kostenfunktion C_g berechnet die Kosten, um L^* in L^+ zu überführen. Diese Überführungskosten werden in der Variable C_g^{old} zur Verfügung gestellt

Generalisierung (Zeile 6-13):

- \mathcal{P} ist die Menge aller möglichen Regelpaare aus L^*
- Das Regelpaar, das beim Merge die geringsten Kosten für die Überführung in die neue Grammatik aufweist, wird in p^* gehalten
- Sind diese Kosten positiv, wird die Generalisierung abgebrochen
- Andernfalls werden die Variablen c^* als Delta-Kosten, C_g^{old} und L^* entsprechen gesetzt
- Sollte die Differenz der Überführungskosten positiv sein, wird die Generalisierung abgebrochen

Algorithmus 38: Generalisieren eines L-Systems mit Gewichtung w_0

```

1  Initialisierung :
2  Regelpaar  $p^* = \emptyset$ 
3   $\mathcal{L}^* = \mathcal{L}^+$ 
4   $C_g^{old} = C_g(\mathcal{L}^* + \{p^*\}, \mathcal{L}^*)$ 
5  Schleife :
6  Finde Regelpaar  $p^*$  mit minimalen Kosten  $C_g(\mathcal{L}^* + \{p_i\}, \mathcal{L}^*), \forall p_i \in \mathcal{P}$ 
7  Wenn  $C_g(\mathcal{L}^* + \{p^*\}, \mathcal{L}^*) \geq 0$ 
8  Breche Schleife ab
9   $c^* = C_g(\mathcal{L}^* + \{p^*\}, \mathcal{L}^*) - C_g^{old}$ 
10  $C_g^{old} = C_g(\mathcal{L}^* + \{p^*\}, \mathcal{L}^*)$ 
11  $\mathcal{L}^* = \mathcal{L}^* + \{p^*\}$ 
12 Wenn  $c^* > 0$ 
13 Breche Schleife ab

```

3.3 Softwarearchitektur

3.3.1 Grundkonzepte

Die Gliederung der Inhalte für die Softwarearchitektur erfolgt nach der arc42-Vorlage [3]

Qualitätsziele

Um die wesentlichen Features des Systems in einem Programm umzusetzen, werden folgende Qualitätsziele definiert, priorisiert (absteigend) und umgesetzt:

- **Funktionalität** durch Umsetzen aller Teilsysteme in Vollständigkeit, Korrektheit, Angemessenheit
- **Interoperabilität** durch Nutzen einer allgemeinen Repräsentation von L-Systemen, damit diese auch in anderen Programmen oder Algorithmen verwendet werden kann
- **Erweiterbarkeit** durch offene Entwurfsmuster (Design Patterns)
- **Modularität** durch Implementierung für effiziente Wartung und Erweiterung
- **Effizienz** durch effiziente Programmierung
- **Attraktivität** durch intuitive Benutzung (Benutzerfreundlichkeit)
- **Plattformunabhängigkeit** durch Verwenden des Java-Frameworks

Kontextabgrenzung

Die Systemgrenzen werden zum Einen durch die Interaktion mit dem Benutzer, zum Anderen durch die Interaktion mit dem Dateisystem des Host-Systems und dem Zugreifen und Lesen der Template-Dateien definiert. Hierbei wird die Erstellung der Basisstruktur als nicht-technische Interaktion und das Einlesen der Dateien als technische Interaktion gesehen.

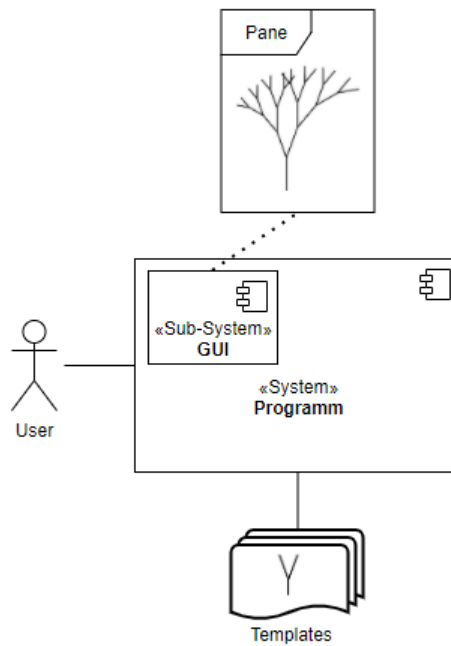


Abbildung 3.1: System und Systemumgebung

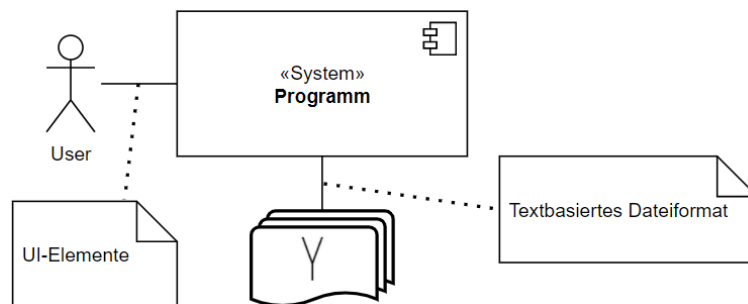


Abbildung 3.2: Interaktion zwischen System und Systemumgebung

Lösungsstrategie

Zur Erreichung der o.g. Qualitätsziele werden folgende Architekturansätze umgesetzt:

Qualitätsziel	Architekturansatz
Funktionalität	Grafische Benutzerschnittstelle Generieren der Baumstruktur Verarbeiten von L-Systemen
Interoperabilität	Durch das Nutzen allgemeingültiger mathematischer Beschreibungen sollen erstellte L-Systeme in Fremdsystemen, wie Online Visualisierer, genutzt werden können
Erweiterbarkeit	Das Nutzen des Pipeline Design Patterns soll das Erweitern des Systems durch Hinzufügen weiterer Teilschritte (Pipes) erleichtern. Trennung der grafischen Oberfläche und der Logik durch Aufbauen des Szenen-graphen über ein XML-Dateiformat
Modularität	Sowohl eine sinnvolle Aufteilung von Funktionalitäten auf Dateien und Software-Pakete, als auch effiziente Datenkapselung und geschlossene Informationskontexte sorgen für Modularität des Programms

Bausteinsicht

Der Benutzer interagiert über das User Interface mit dem Subsystem GUI, das sich mit der Visualisierung, dem Aufbau der Eingabestruktur und dem Anglegen einer internen Baumtopologie beschäftigt. Die Pipeline zum Erzeugen der Ausgabestrukturen beginnt mit dem Inferrieren eines L-Systems aus der benutzerdefinierten Struktur und gibt das erzeugte Ersetzungssystem an die folgenden Komponenten weiter. Hierbei gilt der Pipeline-Kontext, welcher innerhalb der Pipeline an den jeweils nächsten Schritt weitergegeben

und dort aktualisiert wird, als Eingabe der Pipeline. Hat die Estimator-Komponente eine Verteilung über vom Benutzer angelegte Transformationsparameter angelegt, gilt der Kontext als Ausgabe der Pipeline.

Ebene 1

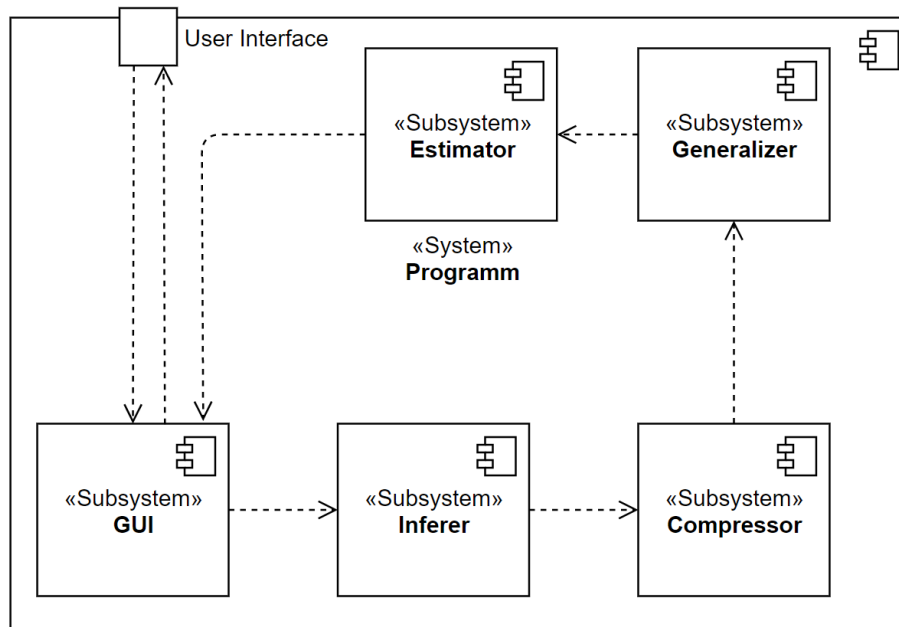


Abbildung 3.3: Subsysteme mit fachlichen Abhängigkeiten

Betrachtet man die GUI-Komponente genauer, setzt sich diese aus vier Subsystemen zusammen. Das Application-Modul beschreibt den Einstiegspunkt für ein JavaFX-Programm. Die grafische Benutzerschnittstelle wird über ein XML-basiertes FXML-Dateischema (Scene graph) mit zugehörigem FXML-Controller (Model) aufgebaut. Der Controller stellt die Logik für eine JavaFX-Oberfläche zur Verfügung. Neben der Erstellung der Basisstruktur, wird die Baumstruktur in einer separaten Komponente umgesetzt.

Ebene 2

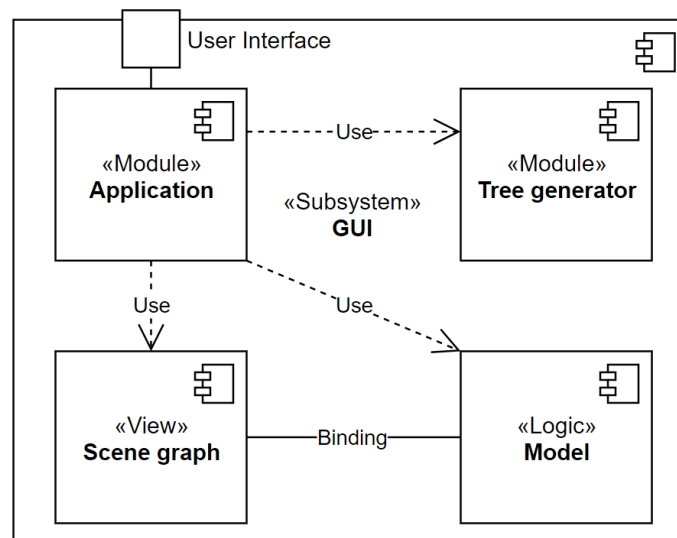


Abbildung 3.4: Subsystem GUI

Laufzeitsicht

Aus der nachfolgenden Abbildung geht hervor, wie sich der Informationsfluss zwischen den einzelnen Subsystemen verhält. Wird der Systemprozess einmal durchlaufen, besteht die Möglichkeit die Ausführung des generalisierten L-Systems mit erneuter Vergabe der Transformationsparameter aus der Häufigkeitsverteilung der Estimator-Komponente zu starten. Der Ablauf des Systems setzt sich wie folgt zusammen:

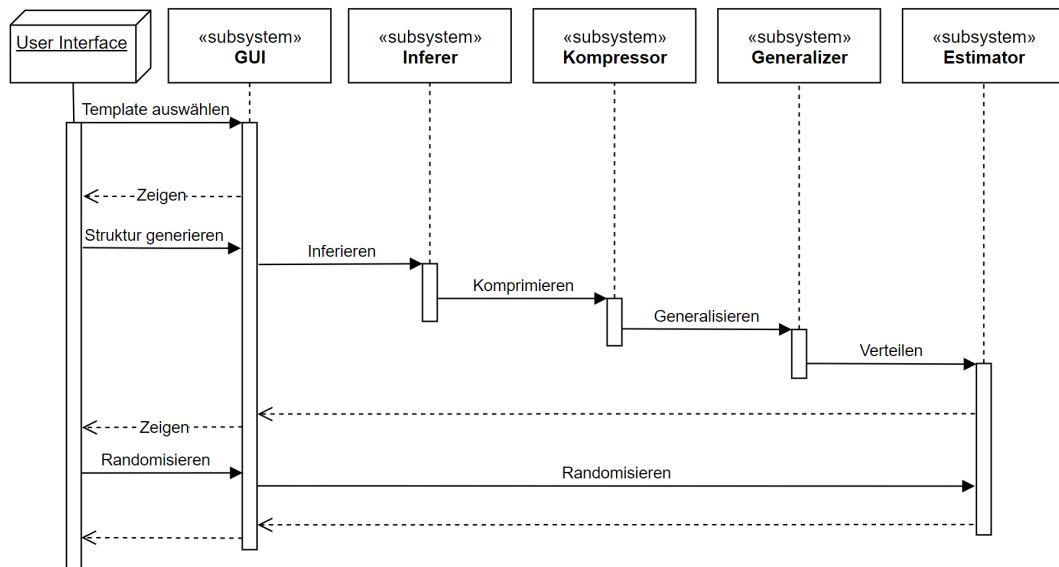


Abbildung 3.5: Laufzeitsicht

Verteilungssicht

Die Ausführung des Programms wird durch ein einfaches Startskript zur Verfügung gestellt. Die folgende Abbildung der Verteilungssicht stellt lediglich die Ausführung auf einer Windows-Maschine dar.

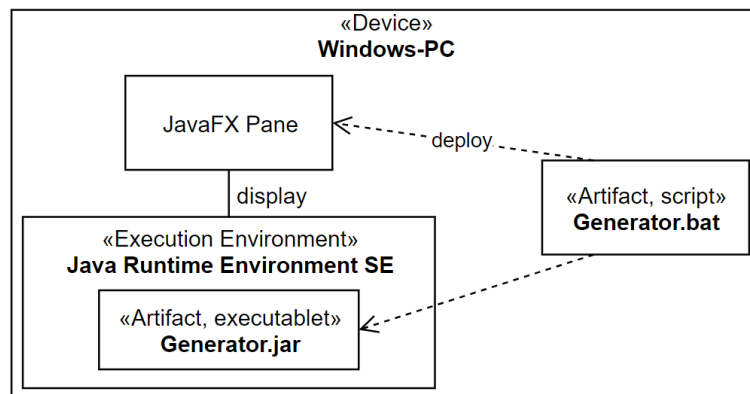


Abbildung 3.6: Infrastruktur Windows-PC

3.3.2 Weitere Konzepte

Testbarkeit

Um eine ausreichende Testabdeckung zu erreichen, werden Klassen als kleinstmöglich zu testende Einheit definiert und durch Komponententests geprüft. Der Name eines Tests setzt sich aus dem Präfix als Name der zu testenden Klasse oder Releases und dem Suffix „Test“ zusammen. Testsubjekte werden als **Blackbox** behandelt, also anhand der Spezifikation getestet.

Bsp.: Klasse *Inferer* mit Komponententest *InfererTest*

Da jedes Release der Implementierung ein funktionsfähiges System beinhaltet, kann auf Integrationstests verzichtet werden. Weiter wird ein Release anhand von funktionalen und nicht-funktionalen Anforderungen getestet. Anhand dieser Systemtests wird geprüft, ob Gesamtspezifikationen umgesetzt worden sind.

Bsp.: Release 2 mit Systemtest *Release2Test*

Zum Schluss der Implementierung wird ein Akzeptanztest durchgeführt.

Validierung

Der Benutzer des Systems nutzt eine grafische Schnittstelle. Somit kann sichergestellt werden, dass dieser keine ungültigen Eingaben tätigt. Werden Template-Dateien unter einem bestimmten Dateipfad nicht gefunden, wird eine *NotFound*-Exception protokolliert und dem Benutzer eine Nachricht über ein Pop-up mitgeteilt. Eine *IllegalArgumentException* wird verzeichnet und ausgegeben, wenn eingelesene Template-Dateien strukturelle Fehler aufweisen.

Fehlerbehandlung

Zur Fehlersuche und -behandlung biete sich eine Protokollierung über Vorgänge, Fehler und Ausnahmen an. Bestimmte Fehler werden dem Benutzer weitergegeben und grafisch angezeigt. Folgende Subsysteme werden in das Programm integriert:

- Ausnahmebehandlung (Exception Handling) und
- Protokollierung (Logging)

Erwartete Exceptions werden jeweils mit eigenen Klassen abgebildet, die von einer entsprechenden Klasse aus der Exception-Hierarchie abgeleitet sind. Das Logging ist statisch und überall im System zugreifbar, um eine einheitliche Protokollierung zu gewährleisten.

3.3.3 Datenstrukturen

TODO: Klassendiagramm aus IntelliJ generieren

4 Implementierung

Zur Umsetzung der vorgestellten Konzepte wird im Folgenden auf Softwarepakete, Technologien, Datenspeicherung, Benutzerinteraktion und Arbeitsablauf der erstellten Software eingegangen. Darüber hinaus wird ein Überblick über genutzte Hardware und einige Implementierungsentscheidungen gegeben. Auf Implementierungsdetails zur Nutzung der vorgestellten Technologien wird nicht eingegangen.

4.1 Projektstruktur

Das Softwareprojekt ist nach der Gradle-Source-Code-Konvention organisiert. Das gesamte Programm befindet sich im Ordner `Generator`, der obligatorische Gradle-Dateien und den Source-Ordner enthält. Sowohl die Quelldateien, also auch die Testdateien sind der Paketstruktur **de.haw** untergeordnet. Die folgende Tabelle gibt einen Überblick über grundlegende Pakete und deren Funktion.

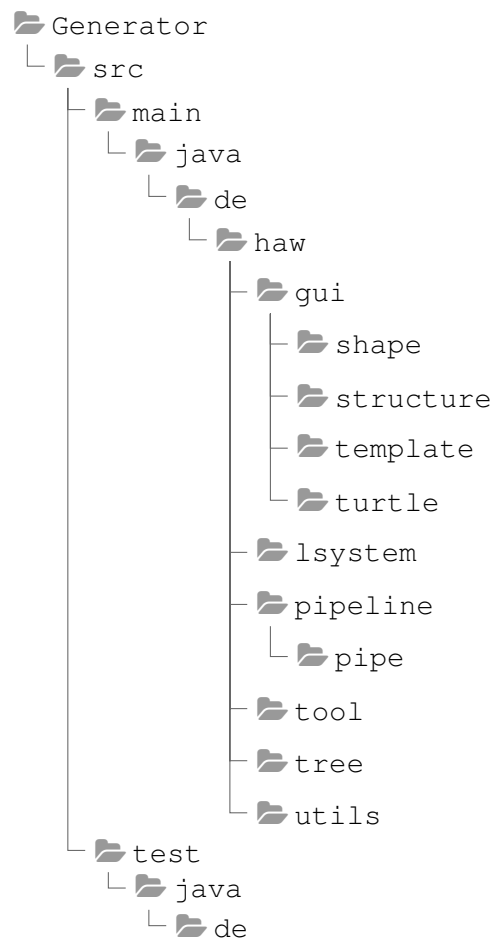


Abbildung 4.1: Softwareprojekt Dateistruktur

Paket	Funktion
<i>gui</i>	Visualisierung des Programms
<i>lsystem</i>	L-System-Repräsentation
<i>pipeline</i>	Umsetzung des Pipeline-Design-Patterns
<i>tool</i>	Methodiken und Algorithmen
<i>tree</i>	Komponenten der Baumstruktur
<i>utils</i>	Hilfskomponenten

Abbildung 4.2: Softwarepakete mit zugehörigen Funktionen

4.2 Technologien

Zur Umsetzung des Softwareprojektes wird eine Java-Anwendung für die Java-Laufzeitumgebung entwickelt. Sie liegt in der Distribution **Amazon Corretto** in der Version 11.0.3_7 vor. Grafische Oberflächen werden mit der JavaFX-Spezifikation von Oracle in der Version 11.0.2 umgesetzt. Zur Automatisierung von Abhängigkeits- und Buildmanagement wird Gradle (Version 6.7) verwendet. Eine Testumgebung, eine Vektorbibliothek, eine Tupelrepräsentation und eine Erweiterung zur mathematischen Standardbibliothek werden über Abhängigkeiten vom Gradle-Framework im Build-Prozess geladen und zur Verfügung gestellt.

Um den test-driven Implementierungsansatz umzusetzen wird JUnit 5 Jupiter als Testumgebung genutzt. Sie setzt sich aus einem Programmierschema und einem Erweiterungsmodell zusammen. Das Jupiter-Projekt liefert zudem die Laufzeitumgebung für Softwaretests. Googles Guava liefert eine Bibliothek mit mathematischen Funktionen. Sie wird benötigt, um eine praktikable Lösung zu Mengenmanipulation nutzen zu können (Bsp. Erstellen von Kombinationspaaren einer Menge). Ausschließlich JavaFX wird außerhalb des Projektes installiert und als Modul im Start-Skript des Programmes hinzugefügt.

Weitere Systeme zur Erstellung des Softwareprojektes sind:

- Versionierung via Git,
- Dot zur Visualisierung von Graphen und
- PlantUML zur Generierung von UML-Diagrammen

4.3 Konzeptumsetzung

Das Skript `Generator.bat`, das zum Starten der Anwendung innerhalb eines Windows-Betriebssystems genutzt werden kann, fügt dem Programm alle externen Module hinzu, die während der Laufzeit genutzt werden, und führt die angegebene jar-Datei aus:

```
1      java --module-path .\javafx-sdk-11.0.2\lib --add-modules  
      ↪     javafx.controls,javafx.fxml,javafx.graphics -jar  
      ↪     Generator-0.1.jar  
2      pause
```

Abbildung 4.3: Startskript Generator.bat

Beim Start der Anwendung findet der Benutzer die grafische Oberfläche vor.

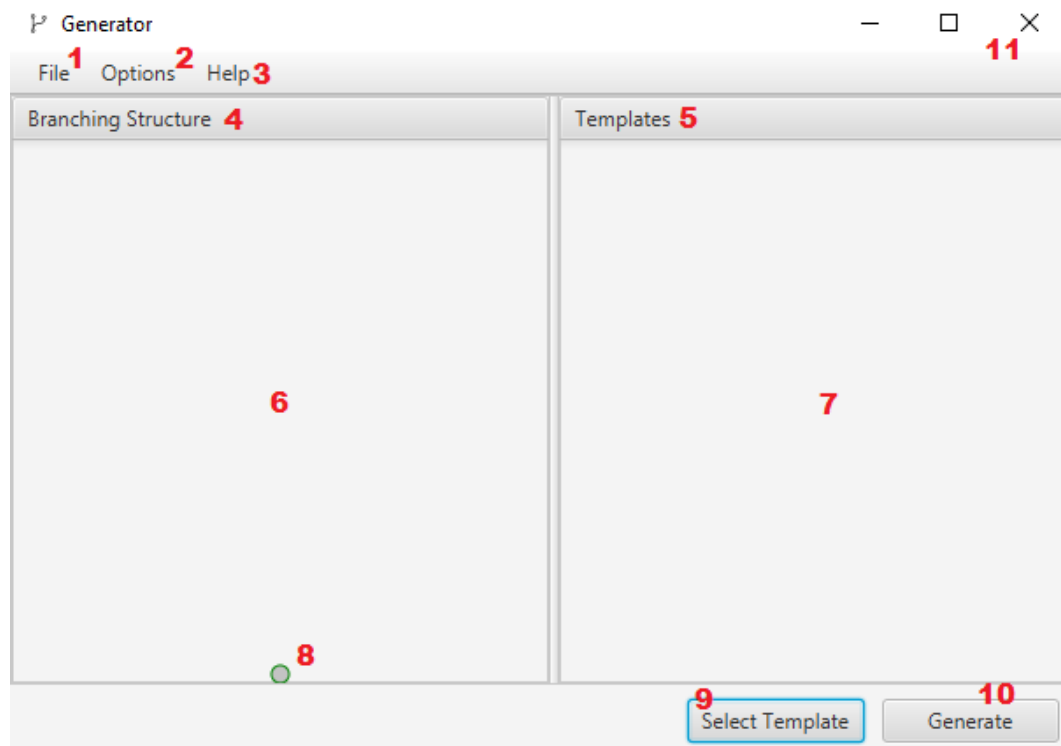


Abbildung 4.4: Programm nach Ausführung des Start-Skripts

Der Menüeintrag **1** bietet Funktionen zum Öffnen des Dateipfades und zum Laden der angelegten Templates-Datei. **2** dient zur Konfiguration der Umgebungsparameter für die Anzahl Iterationen der Ausführung und die Anzahl an Ähnlichkeitsstrukturen, die aus dem resultierenden L-System abgeleitet werden können. Die Gewichtungparameter der vorgestellten Algorithmen können ebenfalls hier eingestellt werden. Die Hilfe (**3**) gibt Informationen zum Projektkontext und zur Benutzung der Anwendung. Die Strukturen **4** und **5** gliedern das Programm in die Ansicht der Verzweigungsstruktur (**6**), die vom

Benutzer angelegt wird, die Übersicht zur Auswahl eines Templates (7) und eine Sicht zur Setzung von Transformationsparametern (7). Der Kreis 8 stellt einen Anker dar, an welchen eine Template-Instanz angehängt werden kann. Sowohl über einen Doppelklick, also auch über einen Button (9) kann ein Template aus der Tempalte-Liste (7) ausgewählt werden. Ist die Verzweigungsstruktur vom Benutzer fertiggestellt, kann die Synthese zur Erstellung der Ähnlichkeitsabbildungen mit 10 erfolgen. 11 schließt die Anwendung.

Visualisierung

Die grafische Oberfläche liegt als MVC-Pattern in Form des Application State (Model), der XML-basierten FXML-Datei (View) und dem JavaFX-Controller vor. Der Arbeitsablauf zur Erstellung der Verzweigungsstruktur kann Schritt für Schritt umgesetzt werden, nachdem die Templates geladen wurden. Dies wird an folgendem Beispiel deutlich gemacht. Die Parameter **Number of iterations**, **Number of generations**, **Rule application ratio** und **Merge application ratio** werden auf die Werte 8, 5, 0.5 und 0.5 gesetzt (Standarteinstellung).

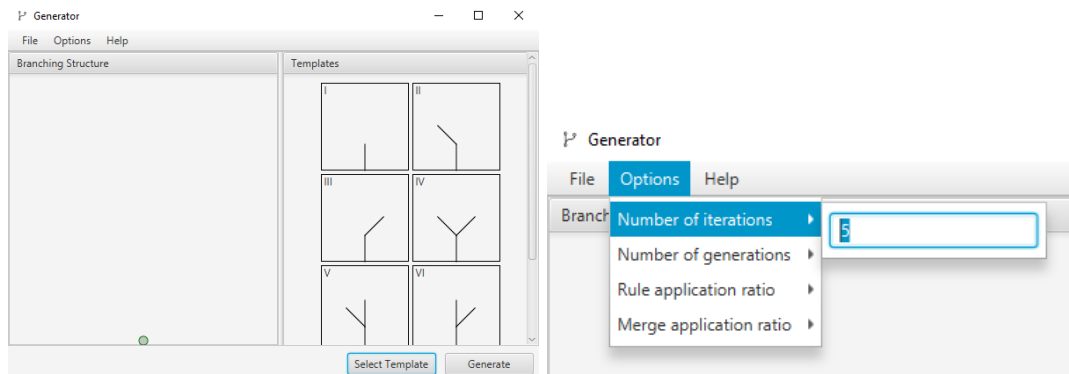


Abbildung 4.5: Erster Anker ist vorselektiert & gesetzte Parameter

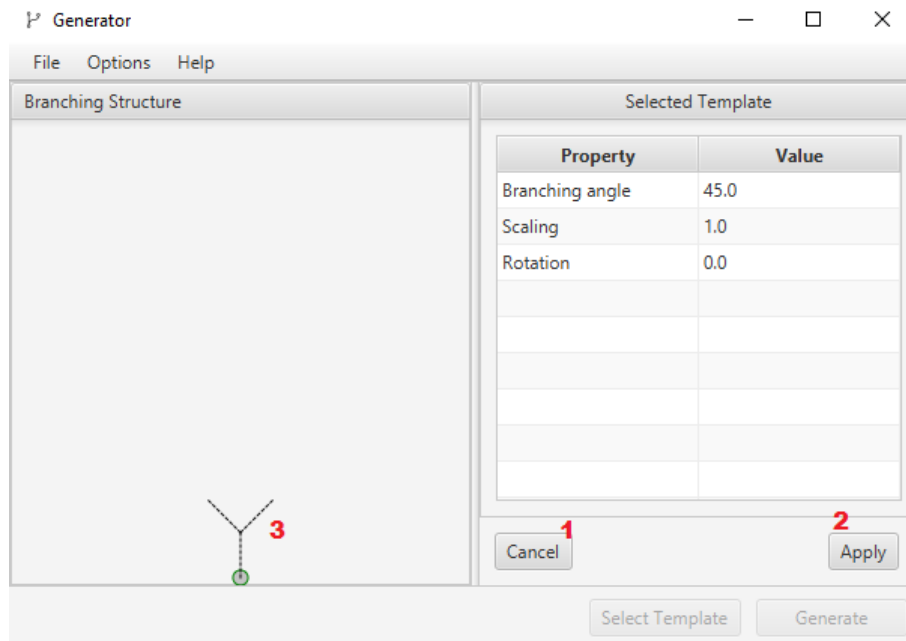


Abbildung 4.6: Auswahl des ersten Templates

Nachdem ein Template ausgewählt wurde, können die Transformationsparameter gesetzt (Doppelklick) werden und die Auswahl rückgängig gemacht (**1**) oder bestätigt (**2**) werden. **3** zeigt einen Entwurf der Template-Instanz, die mit den aktuellen Transformationsparametern angepasst wurde. Mit der Bestätigung wird die Template-Instanz endgültig der Basisstruktur hinzugefügt und der Benutzer gelangt wieder in die Übersicht der zur Verfügung stehenden Templates.

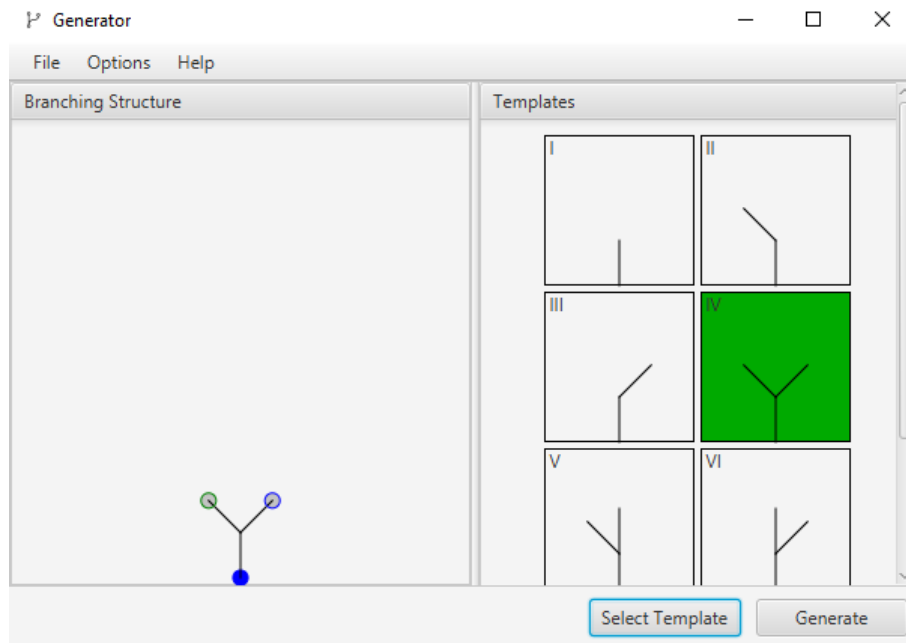


Abbildung 4.7: Der Verzweigungsstruktur hinzugefügte Template-Instanz

Dieser Vorgang wiederholt sich, bis der Benutzer die Struktur fertiggestellt hat.

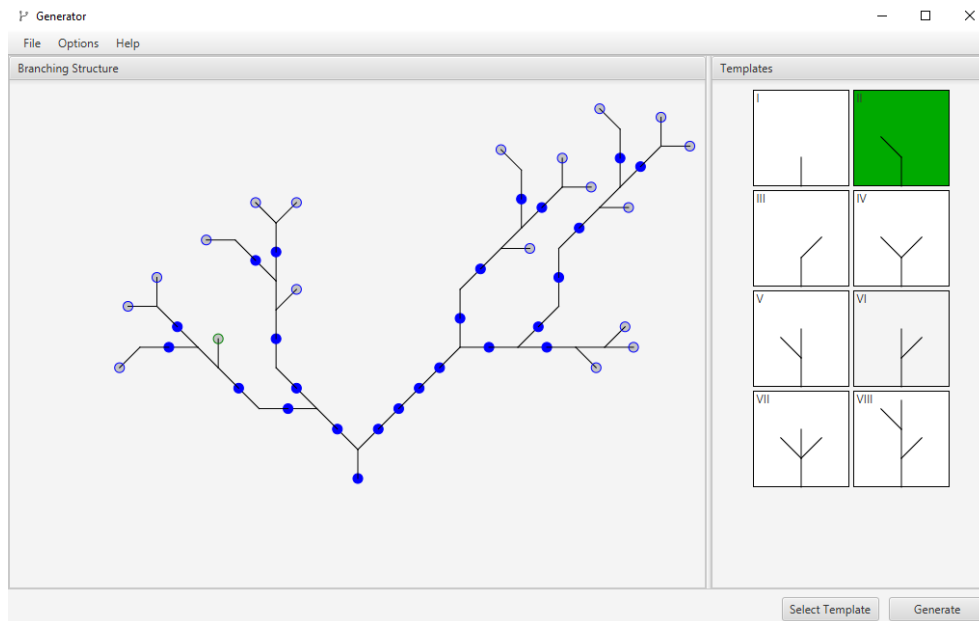


Abbildung 4.8: Vom Benutzer fertiggestellte Verzweigungsstruktur

Baumtopologie

Um eine interne Baumtopologie aufzubauen wird die Klasse `TreeNode` verwendet, um eine iterierbare Baumstruktur aufzubauen:

```
/**
 * Iterable tree node structure containing a payload and children
 * as tree nodes.
 * Every node represents a whole tree
 * @param <T> Tree node content
 */
public class TreeNode<T> implements Iterable<TreeNode<T>> {
    // Payload
    private T data;
    // Children
    protected List<TreeNode<T>> children;
    ...
    @Override
    public Iterator<TreeNode<T>> iterator() {
        return new TreeNodeIterator<>(this);
    }
    ...
}
```

`TreeNode` implementiert die `Iterable`-Schnittstelle, was ein Überladen der `iterator`-Funktion möglich macht. Sie stellt einen Iterator des Baumes zur Verfügung. Dieser Iterator (`TreeNodeIterator`) implementiert die `Iterator`-Schnittstelle und definiert so die Funktionen `hasNext` und `next`.

```
/**
 * Tree node iterator for iterating nodes of a tree
 * @param <T> Payload of the tree nodes
 */
public class TreeNodeIterator<T> implements Iterator<TreeNode<T>>{
    ...
    @Override
    public boolean hasNext() {
        return !queue.isEmpty();
    }

    @Override
    public TreeNode<T> next() {
        if (queue.isEmpty()) return null;
        var node = queue.pop();
    }
}
```

```

        queue.addAll(node.getChildren());
        return node;
    }
}

```

Die Funktion `next()` zeigt, wie eine Iteration des Baumes als Breitensuche implementiert ist. Diese wird beim Inferrieren eines L-Systems aus der Baumstruktur benötigt. Das folgende Beispiel zeigt eine erstellte Baumstruktur (links) und deren erstellte Baumtopologie (rechts).

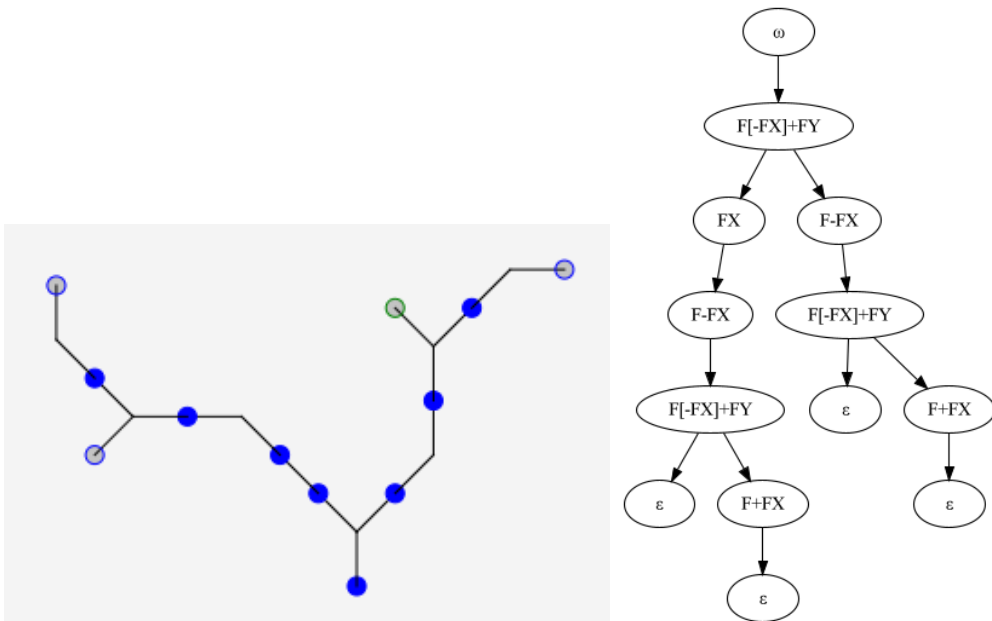


Abbildung 4.10: Verzweigungsstruktur & zugehörige Baumstruktur

Prozesspipeline

4.4 Entscheidungen

Änderbarkeit

Risiken

Qualitätsmerkmale

Alternativen

Aufwand der Implementierung

5 Evaluierung

...

5.1 Testumgebung

...

5.2 Beobachtungen & Ergebnisse

...

5.3 Diskussion und Bewertung

...

6 Ausblick

...

Literaturverzeichnis

- [1] ALHALAWANI, Sawsan ; YANG, Yong-Liang ; LIU, Han ; MITRA, Niloy: Interactive Facades Analysis and Synthesis of Semi-Regular Facades. In: *Computer Graphics Forum* 32 (2013), 05
- [2] ALIAGA, Daniel G. ; DEMIR, undefinedlke ; BENES, Bedrich ; WAND, Michael: Inverse Procedural Modeling of 3D Models for Virtual Worlds. In: *ACM SIGGRAPH 2016 Courses*. New York, NY, USA : Association for Computing Machinery, 2016 (SIGGRAPH '16). – URL <https://doi.org/10.1145/2897826.2927323>. – ISBN 9781450342896
- [3] HRUSCHKA, Dr. P. ; STARKE, Dr. G.: *arc42 Softwarearchitektur-Template*. <https://arc42.de/template/>. – Pragmatisches Muster für die Erstellung, Dokumentation und Kommunikation von Software- und Systemarchitekturen
- [4] BENES, Bedrich ; STAVA, Ondrej ; MECH, Radomir ; MILLER, Gavin: Guided Procedural Modeling. In: *Computer Graphics Forum* 30 (2011), 05, S. 325–334
- [5] CHOMSKY, N.: Three models for the description of language. In: *IRE Transactions on Information Theory* 2 (1956), Nr. 3, S. 113–124
- [6] DAELEMANS, Walter: Colin de la Higuera: Grammatical inference: learning automata and grammars: Cambridge University Press, 2010, iv + 417 pages. In: *Machine Translation* 24 (2010), 12, S. 291–293
- [7] DEUSSEN, Oliver ; LINTERMANN, Bernd: *Digital Design of Nature: Computer Generated Plants and Organics*. 1st. Springer Publishing Company, Incorporated, 2010. – ISBN 3642073638
- [8] *Git - the stupid content tracker*. <https://git-scm.com/>. – Git is a member of Software Freedom Conservancy
- [9] *Github*. <https://github.com/adrian-helberg/bachelor>. – GitHub, Inc. (2020)

- [10] *Gradle Build Tool*. <https://gradle.org/>. – Gradle Inc. 2020
- [11] GUO, Jianwei ; JIANG, Haiyong ; BENES, Bedrich ; DEUSSEN, Oliver ; ZHANG, Xiaopeng ; LISCHINSKI, Dani ; HUANG, Hui: Inverse Procedural Modeling of Branching Structures by Inferring L-Systems. In: *ACM Trans. Graph.* 39 (2020), Juni, Nr. 5. – URL <https://doi.org/10.1145/3394105>. – ISSN 0730-0301
- [12] HIGUERA, Colin De la: *Grammatical Inference: Learning Automata and Grammars*. USA : Cambridge University Press, 2010. – ISBN 0521763169
- [13] *Jetbrains IntelliJ IDEA*. <https://www.jetbrains.com/idea/>. – Capable and Ergonomic IDE for JVM
- [14] LINDENMAYER, A: Mathematical models for cellular interactions in development. I. Filaments with one-sided inputs. In: *Journal of theoretical biology* 18 (1968), March, Nr. 3, S. 280–299. – URL [https://doi.org/10.1016/0022-5193\(68\)90079-9](https://doi.org/10.1016/0022-5193(68)90079-9). – ISSN 0022-5193
- [15] MARTINOVIC, Andelo ; VAN GOOL, Luc: Bayesian Grammar Learning for Inverse Procedural Modeling, 06 2013, S. 201–208
- [16] MCQUILLAN, Ian ; BERNARD, Jason ; PRUSINKIEWICZ, Przemyslaw: *Algorithms for Inferring Context-Sensitive L-Systems*. S. 117–130, 01 2018. – ISBN 978-3-319-92434-2
- [17] MERRELL, Paul ; SCHKUFZA, Eric ; LI, Zeyang ; AGRAWALA, Maneesh ; KOLTUN, Vladlen: *Interactive Furniture Layout Using Interior Design Guidelines*. In: *ACM SIGGRAPH 2011 Papers*. New York, NY, USA : Association for Computing Machinery, 2011. – URL <https://doi.org/10.1145/1964921.1964982>. – ISBN 9781450309431
- [18] MÜLLER, Pascal ; WONKA, Peter ; HAEGLER, Simon ; ULMER, Andreas ; VAN GOOL, Luc: Procedural Modeling of Buildings. In: *ACM Trans. Graph.* 25 (2006), 07, S. 614–623
- [19] MULLER, M. M. ; TICHY, W. F.: Case study: extreme programming in a university environment. In: *Proceedings of the 23rd International Conference on Software Engineering. ICSE 2001*, 2001, S. 537–544
- [20] NISHIDA, Gen ; GARCIA-DORADO, Ignacio ; ALIAGA, Daniel G. ; BENES, Bedrich ; BOUSSEAU, Adrien: Interactive Sketching of Urban Procedural Models. In: *ACM*

- Trans. Graph.* 35 (2016), Juli, Nr. 4. – URL <https://doi.org/10.1145/2897824.2925951>. – ISSN 0730-0301
- [21] PARISH, Yoav ; MÜLLER, Pascal: Procedural Modeling of Cities, 08 2001, S. 301–308
- [22] PRUSINKIEWICZ, P: Graphical Applications of L-Systems. In: *Proceedings on Graphics Interface '86/Vision Interface '86*. CAN : Canadian Information Processing Society, 1986, S. 247–253
- [23] PRUSINKIEWICZ, P. ; LINDENMAYER, Aristid: *The Algorithmic Beauty of Plants*. Berlin, Heidelberg : Springer-Verlag, 1990. – ISBN 0387972978
- [24] SMELIK, Ruben M. ; TUTENEL, Tim ; BIDARRA, Rafael ; BENES, Bedrich: A Survey on Procedural Modelling for Virtual Worlds. In: *Comput. Graph. Forum* 33 (2014), September, Nr. 6, S. 31–50. – URL <https://doi.org/10.1111/cgf.12276>. – ISSN 0167-7055
- [25] STAVA, O. ; PIRK, S. ; KRATT, J. ; CHEN, B. ; MUNDEFINCH, R. ; DEUSSEN, O. ; BENES, B.: Inverse Procedural Modelling of Trees. In: *Comput. Graph. Forum* 33 (2014), September, Nr. 6, S. 118–131. – URL <https://doi.org/10.1111/cgf.12282>. – ISSN 0167-7055
- [26] STAVA, Ondrej ; BENES, Bedrich ; MECH, Radomir ; ALIAGA, Daniel ; KRISTOF, Peter: Inverse Procedural Modeling by Automatic Generation of L-systems. In: *Computer Graphics Forum* 29 (2010), 05, S. 1467–8659
- [27] TALTON, Jerry ; LOU, Yu ; DUKE, Jared ; LESSER, Steve ; MECH, Radomir ; KOLTUN, Vladlen: Metropolis Procedural Modeling. In: *ACM Transactions on Graphics* 30 (2011), 04
- [28] TALTON, Jerry ; YANG, Lingfeng ; KUMAR, Ranjitha ; LIM, Maxine ; GOODMAN, Noah ; MECH, Radomir: Learning design patterns with Bayesian grammar induction. In: *UIST'12 - Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (2012), 10. ISBN 978-1-4503-1580-7
- [29] *Trello Board*. <https://trello.com/>. – Trello Board, 2020 Atlassian
- [30] VAZQUEZ, Pablo ; SIDDI, Francesco: *About Blender*. <https://www.blender.org/about/>. 2017. – Blender Foundation (2002)
- [31] XIAO, Jianxiong ; FANG, Tian ; ZHAO, Peng ; OFEK, Eyal ; QUAN, Long: Image-based Façade Modeling. In: *ACM Transactions on Graphics* 27 (2008), 12, S. 161

- [32] ZHANG, Song-Hai ; ZHANG, Shao-Kui ; LIANG, Yuan ; HALL, Peter: A Survey of 3D Indoor Scene Synthesis. In: *Journal of Computer Science and Technology* 34 (2019), 05, S. 594–608

A Anhang

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Glossar

HAW Hamburg Die HAW Hamburg ist die formalige Fachhochschule am Berliner Tor.

Erklärung zur selbstständigen Bearbeitung einer Abschlussarbeit

Gemäß der Allgemeinen Prüfungs- und Studienordnung ist zusammen mit der Abschlussarbeit eine schriftliche Erklärung abzugeben, in der der Studierende bestätigt, dass die Abschlussarbeit „— bei einer Gruppenarbeit die entsprechend gekennzeichneten Teile der Arbeit [(§ 18 Abs. 1 APSO-TI-BM bzw. § 21 Abs. 1 APSO-INGI)] — ohne fremde Hilfe selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt wurden. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quellen kenntlich zu machen.“

Quelle: § 16 Abs. 5 APSO-TI-BM bzw. § 15 Abs. 6 APSO-INGI

Erklärung zur selbstständigen Bearbeitung der Arbeit

Hiermit versichere ich,

Name: _____

Vorname: _____

dass ich die vorliegende Bachelorarbeit – bzw. bei einer Gruppenarbeit die entsprechend gekennzeichneten Teile der Arbeit – mit dem Thema:

Template-basierte Synthese von Verzweigungsstrukturen mittels L-Systemen

ohne fremde Hilfe selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quellen kenntlich gemacht.

_____	_____	_____
Ort	Datum	Unterschrift im Original