UNIVERSITÄT KOBLENZ · LANDAU

Megamodel-driven Traceability Recovery & Exploration of Correspondence & Conformance Links

Bachelorarbeit

zur Erlangung des Grades eines Bachelor of Science
im Studiengang Informatik

vorgelegt von

Maximilian Meffert

Erstgutachter: Prof. Dr. Ralf Lämmel
Institut für Informatik

Zweitgutachter: Msc. Johannes Härtel
Institut für Informatik

Koblenz, im Juni 2017

Erklärung

Hiermit bestätige ich, dass die vorliegende Arbeit von mir selbständig verfasst wurde und ich keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internet–Quellen – benutzt habe und die Arbeit von mir vorher nicht in einem anderen Prüfungsverfahren eingereicht wurde. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium (CD-Rom).

	Ja	Nein
Mit der Einstellung der Arbeit in die Bibliothek bin ich einverstanden.	☐	☐
Der Veröffentlichung dieser Arbeit im Internet stimme ich zu.	☐	☐

....

(Ort, Datum) (Maximilian Meffert)

Zusammenfassung

TBD.

Abstract

TBD.

Contents

List of Figures

Chapter 1

Introduction

TBD.

Chapter 2

Related Work

TBD.

Chapter 3

Background

This section summarizes the necessary background topics of the thesis. Each topic is introduced independently, interrelation is done during synthesis of hypotheses for this thesis in chapter 4.

3.1 Relations

This section introduces the necessary aspects of mathematical relations for this thesis. The concept of relations is a generalization of semantic dependencies between two or more mathematical objects. This section is based on [1].

Relations are based on set-theory. We also introduce the necessary constructs of set-theory in order to clarify terminology and notation. A set is a collection of well distinguishable mathematical objects. Objects in a set are called elements of the set. A set does not contain two or more identical elements. The notation $x \in X$ denotes that x is an element of the set X.

Definition 1 (Inclusion) *Let X and Y be a sets. Y includes X if and only if:*

$$X \subset Y :\Leftrightarrow \forall x[x \in X \Rightarrow x \in Y] \tag{3.1}$$

Then X is called subset *of Y and Y is called superset of X.*

Definition 2 (Power-Set) *Let X be a set, then the* power-set *of X is defined as:*

$$\mathcal{P}(X) := \{Y | Y \subset X\} \tag{3.2}$$

From now on we assume that there is a universal set, called *universe*, containing all elements of interest.

Definition 3 (Union) *Let U be a universe and $X, Y \in \mathcal{P}(U)$.*

$$X \cup Y := \{x | x \in X \vee x \in Y\} \tag{3.3}$$

$X \cup Y$ is called union *of X and Y.*

Definition 4 (Intersection) *Let U be a universe and $X, Y \in \mathcal{P}(U)$.*

$$X \cap Y := \{x | x \in X \wedge x \in Y\} \tag{3.4}$$

$X \cap Y$ is called intersection *of X and Y.*

The definition of inclusion provides an order for power-sets. So we may compare two sets A and B in the sense of *smaller* and *larger*, i.e.:

$$A \text{ is smaller than } B \Leftrightarrow A \subset B \tag{3.5}$$

$$A \text{ is larger than } B \Leftrightarrow A \supset B \tag{3.6}$$

$$A \text{ is the smallest subset of } B \Leftrightarrow \forall C \in \mathcal{P}(B) : A \subset C \tag{3.7}$$

$$A \text{ is the largest subset of } B \Leftrightarrow \forall C \in \mathcal{P}(B) : A \supset C \tag{3.8}$$

Definition 5 (Upper & Lower Bound) *Let U be a universe, $X \in \mathcal{P}(U)$ be sets in the universe and $A \subset \mathcal{P}(U)$ subsets of the universe.*

$$X \textit{ is an } \text{upper bound} \textit{ for } A :\Leftrightarrow \forall Y \in A : Y \subset X \tag{3.9}$$

$$X \textit{ is a } \text{lower bound} \textit{ for } A :\Leftrightarrow \forall Y \in A : Y \supset X \tag{3.10}$$

We also define:

- $\mathbf{U}_A := \{B \in \mathcal{P}(U) | \forall Y \in A : Y \subset B\}$ *as the set of all upper bounds of A*
- $\mathbf{L}_A := \{B \in \mathcal{P}(U) | \forall Y \in A : Y \supset B\}$ *as the set of all lower bounds of A*

Definition 6 (Supremum & Infimum) *Let U be a universe, $X \in \mathcal{P}(U)$ be sets in the universe and $A \subset \mathcal{P}(U)$ subsets of the universe.*

$$X = \sup A \Leftrightarrow X \textit{ is } \text{supremum} \textit{ for} A \quad :\Leftrightarrow X \in \mathbf{U}_A \wedge \forall Y \in \mathbf{U}_A : X \subset Y \tag{3.11}$$

$$X = \inf A \Leftrightarrow X \textit{ is } \text{infimum} \textit{ for} A \quad :\Leftrightarrow X \in \mathbf{L}_A \wedge \forall Y \in \mathbf{L}_A : X \supset Y \tag{3.12}$$

For arbitrary subsets $A \subset \mathcal{P}(U)$ supremum and infimum are:

$$\sup A := \bigcup_{X \in A} X \tag{3.13}$$

$$\inf A := \bigcap_{X \in A} X \tag{3.14}$$

Definition 7 (Cartesian Product) *Let U be a universe and $X_n \in \mathcal{P}(U)$ sets with $i = 1...n, n \in \mathbb{N}$, then:*

$$X_1 \times ... \times X_n := \{(x_1, ..., x_n)\} \tag{3.15}$$

is called Cartesian product*.*

Definition 8 (Relation) *A relation is a subset of a Cartesian product:*

$$R \subset X_1 \times ... \times X_n \tag{3.16}$$

The relation of only two sets is called binary relation*:*

An arbitrary relation $R \subset A \times B$ is called *homogeneous* if $A = B$, otherwise it is called *heterogeneous*. However, an arbitrary relation $R \subset A \times B$ is also homogeneous in the sense of $R \subset A \times B \subset (A \cup B) \times (A \cup B)$.

Definition 9 (Properties of Relations)

$$\text{bijective} :\Leftrightarrow \forall b \in B \exists! a \in A : (a, b) \in R \tag{3.17}$$

$$\text{function} :\Leftrightarrow \forall a \in A \exists! b \in B : (a, b) \in R \tag{3.18}$$

$$\text{reflexive} :\Leftrightarrow \forall a \in A : (a, a) \in R \tag{3.19}$$

$$\text{irreflexive} :\Leftrightarrow \forall a \in A : (a, a) \notin R \tag{3.20}$$

$$\text{transitive} :\Leftrightarrow \forall a, b, c \in A : (a, b) \in R \wedge (b, c) \in R \Rightarrow (a, c) \in R \tag{3.21}$$

$$\text{intransitive} :\Leftrightarrow \forall a, b, c \in A : (a, b) \in R \wedge (b, c) \in R \Rightarrow (a, c) \notin R \tag{3.22}$$

$$\text{symmetric} :\Leftrightarrow \forall a, b \in A : (a, b) \in R \Rightarrow (b, a) \in R \tag{3.23}$$

$$\text{asymmetric} :\Leftrightarrow \forall a, b \in A : (a, b) \in R \Rightarrow (b, a) \notin R \tag{3.24}$$

$$\text{antisymmetric} :\Leftrightarrow \forall a, b \in A : (a, b) \in R \wedge (b, a) \in R \Rightarrow a = b \tag{3.25}$$

Definition 10 (Composition of Binary Relations) *The composition $S \circ R \subseteq A \times D$ of two binary relations $R \subseteq A \times B$ and $S \subseteq C \times D$ is defined as:*

$$S \circ R := \{(a, d) \in A \times D | \exists x \in B \cap C : (a, x) \in R \wedge (x, d) \in S\} \tag{3.26}$$

Definition 11 (Identity Relation) *The relation I over a set A with:*

$$I := \{(a, b) \in A \times A | a = b\} = \{(a, a) | a \in A\} \subseteq A \times A \tag{3.27}$$

is called identity relation.

The n-th power R^n of a relation R with $n \in \mathbb{N}$ is the product of its composition n-times with itself:

$$R^n = \underbrace{R \circ ... \circ R}_{n} \tag{3.28}$$

and can be recursively defined with:

$$R^0 := I \tag{3.29}$$

$$R^n := R \circ R^{n-1} \tag{3.30}$$

3.1.1 Closures

For a relation R over a set A its closure:

- $R \cup I$ is called *reflexive closure*
- $R^+ := R^1 \cup R^2 \cup R^3 \cup ...$ is called *transitive closure*

$$\begin{aligned}
(a, b) \in R^+ &\Leftrightarrow (a, b) \in R^1 \vee (a, b) \in R^2 \vee (a, b) \in R^3 \vee ... && (3.31)\\
&\Leftrightarrow \exists n \in \mathbb{N} : n \geq 1 \wedge (a, b) \in R^n && (3.32)\\
&\Leftrightarrow \exists n \in \mathbb{N} : n \geq 1 \wedge (a, b) \in R \circ R^{n-1} && (3.33)\\
&\Leftrightarrow \exists n \in \mathbb{N} : n \geq 1 \wedge [\exists x \in A : (a, x) \in R \wedge (x, b) \in R^{n-1}] && (3.34)\\
&\Leftrightarrow \exists n \in \mathbb{N} \exists x \in A : n \geq 1 \wedge (a, x) \in R \wedge (x, b) \in R^{n-1} && (3.35)
\end{aligned}$$

- $R^* := R^+ \cup I$ is called *reflexive-transitive closure*

$$(a,b) \in R^* \Leftrightarrow (a,b) \in R^+ \vee (a,b) \in I \tag{3.36}$$

$$\Leftrightarrow (a,b) \in R^+ \vee a = b \tag{3.37}$$

3.2 Formal Languages & Grammars

3.2.1 Context-Free Languages & Grammars

3.3 Traceability

[2] TBD.

3.3.1 Traceability Relationship

3.3.2 Traceability Link

3.3.3 Traceability Recovery

3.3.4 Traceability Exploration

3.4 Megamodeling

TBD.

3.4.1 MegaL

3.4.1.1 MegaL/Xtext

3.5 Ontologies

TBD.

3.6 Mereology

$$x \ \mathsf{partOf}\ x \qquad \text{(Reflexivity)} \tag{3.38}$$

$$x \ \mathsf{partOf}\ y \wedge y \ \mathsf{partOf}\ x \Rightarrow x = y \qquad \text{(Antisymmetry)} \tag{3.39}$$

$$x \ \mathsf{partOf}\ y \wedge y \ \mathsf{partOf}\ z \Rightarrow x \ \mathsf{partOf}\ z \qquad \text{(Transitivity)} \tag{3.40}$$

TBD.

3.7 Program Analysis

TBD.

3.8 XML Data Binding

TBD.

3.8.1 Java Architecture for XML Binding (JAXB)

3.9 Object Relational Mapping

TBD.

3.9.1 Java Persistence API (JPA)

3.9.2 Hibernate

3.10 Another Tool For Language Recognition (ANTLR)

Chapter 4

Hypotheses

TBD.

4.1 Fragments

4.2 Correspondence

4.3 Conformance

Chapter 5

Methodology

TBD.

5.1 The 101companies Human Resource Management System

Description of the 101companies Human Resource Management System

5.2 Link Proper Part Ratio

The ratio between all proper parts of two artifacts and the proper parts of the same artifacts in a relationship.

$$\pi_{R,A_1,A_2} = \frac{|\{(p_1, p_2) \in R : p_1 \text{ properPartOf } A_1 \wedge p_2 \text{ properPartOf } A_2\}|}{|\{p : p \text{ properPartOf } A_1 \vee p \text{ properPartOf } A_2\}|}$$

Chapter 6

Requirements

TBD.

R1 asdf

R2 asdf

R3 asdf

R4 asdf

Chapter 7

Design

TBD.

Chapter 8

Implementation

TBD.

8.1 Context-Free Grammar Fragmentation

8.2 Name Correspondence Heuristic

Heuristics are quick and "simple" methods for finding good approximate solutions for complex problems. The Name Correspondence Heuristic determines correspondence between artifacts simply by finding similarities of names in those artifacts.

TBD.

Chapter 9

Results

TBD.

Chapter 10

Conclusion

TBD.

Bibliography

[1] Gunther Schmidt and Thomas Ströhlein. *Relationen und Graphen*. Mathematik für Informatiker. Springer, 1989. ISBN: 3-540-50304-8.

[2] Stefan Winkler and Jens Pilgrim. "A Survey of Traceability in Requirements Engineering and Model-driven Development". In: *Softw. Syst. Model.* 9.4 (Sept. 2010), pp. 529–565. ISSN: 1619-1366. DOI: 10.1007/s10270-009-0145-0. URL: http://dx.doi.org/10.1007/s10270-009-0145-0.