

Brandenburgische Technische Universität Cottbus-Senftenberg
Institut für Informatik
Fachgebiet Praktische Informatik/Softwaresystemtechnik

Masterarbeit



Brandenburgische
Technische Universität
Cottbus - Senftenberg

Testdesign Automatisierung von GraphQL mittels graphspezifischen Domänenwissen

Testdesign Automation for GraphQL using graphspecific domain knowledge

Tom Lorenz

MatrikelNr.: 3711679

Studiengang: Informatik M.Sc

Datum der Themenausgabe: (hier einfügen)

Datum der Abgabe: (hier einfügen)

1. Betreuer: Prof. Dr. rer. nat. Leen Lambers
2. Betreuer: M.Sc Lucas Sakizloglou

Eidesstattliche Erklärung

Der Verfasser erklärt, dass er die vorliegende Arbeit selbständig, ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt hat. Die aus fremden Quellen (einschließlich elektronischer Quellen) direkt oder indirekt übernommenen Gedanken sind ausnahmslos als solche kenntlich gemacht. Wörtlich und inhaltlich verwendete Quellen wurden entsprechend den anerkannten Regeln wissenschaftlichen Arbeitens zitiert. Die Arbeit ist nicht in gleicher oder vergleichbarer Form (auch nicht auszugsweise) im Rahmen einer anderen Prüfung bei einer anderen Hochschule vorgelegt oder publiziert worden. Der Verfasser erklärt sich zudem damit einverstanden, dass die Arbeit mit Hilfe eines Plagiatserkennungsdienstes auf enthaltene Plagiate überprüft wird.

.....
Ort, Datum

.....
Unterschrift

Inhaltsverzeichnis

Table of Contents	ii
0.1 Abstract	1
0.2 Einleitung	2
0.2.1 Motivation	2
0.2.2 Umsetzung	3
0.3 Glossar	8
Literaturverzeichnis	9

0.1 Abstract

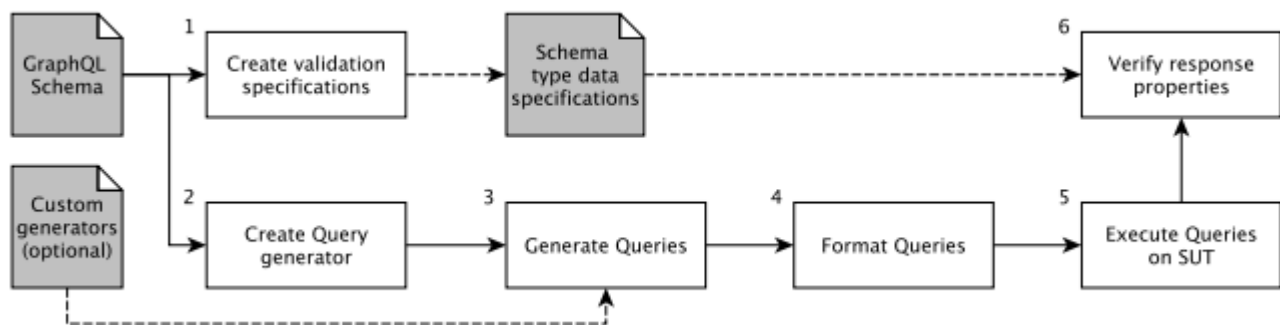
Mit zunehmender Popularität von GraphQL ist es wichtig auch die Qualität von GraphQL-API's zu testen. Aktuell gibt es aber noch eine Lücke an Testtools für GraphQL-Apis. Im Paper "Automatic Property-based Testing of GraphQL APIs" (Quelle hinzufügen) wurde sich mit einem automatischen Testverfahren für GraphQL-API's beschäftigt. Allerdings bietet diese Arbeit ein Verbesserungspotential, welches in dieser Arbeit untersucht und implementiert werden soll. Im konkreten handelt es sich bei dem automatischen Testverfahren um ein Verfahren, dass ein GraphQL-Schema aufgrund seiner Eigenschaften aufspaltet und Testet. Hierbei wird zwar Rücksicht auf die Graphstruktur genommen allerdings werden spezifische Grapheigenschaften nicht ausgenutzt um die Tests zu verbessern. Ziel dieser Arbeit ist es, das Domänenwissen für Graphen zu nutzen um eine automatische Testgenerierung zu verbessern und algorithmisch beweisbar eine ideale Abdeckung mit Tests zu erreichen.

0.2 Einleitung

In diesem Kapitel wird an das Thema und die Motivation dieser Arbeit herangeführt. Außerdem wird definiert, welche Ziele diese Arbeit erreichen soll und eine grobe Übersicht über die Kapitelstruktur gegeben.

0.2.1 Motivation

Mit einer steigenden Nutzung von GraphQL wird es immer wichtiger, geeignete Tests für GraphQL-API's zu entwickeln damit eine gute Softwarequalität sichergestellt werden kann. Idealerweise können diese Testtools solche API's automatisch testen, so wie es für REST-API's schon umgesetzt wurde. Die Struktur von GraphQL erlaubt allerdings zyklische Strukturen und ermöglicht somit ein potenziell unendlich großen Testraum. In einem Paper "Automatic Property-based Testing of GraphQL-API's" (hier Quelle) wurde versucht ein solches automatisches Testtool schon umzusetzen. Allerdings hat dieses Paper zwei große Verbesserungspunkte. Einerseits die Art, wie die Tests generiert werden, andererseits die Auswertung der Tests. Der allgemeine Ablauf des bestehenden Tools ist wie folgt:



Verbesserungen in dieser Arbeit sind insbesondere in den Punkten 2,3 und 6 geplant.

Create Query Generator (Punkt 2) Kapitel Testgenerierung

Generate Queries (Punkt 3) Kapitel Testgenerierung

Verify response properties (Punkt 6) Kapitel Testauswertung

Testgenerierung

Bei der Testgenerierung wurde mit den zyklischen Strukturen in GraphQL Schemas so umgegangen, dass eine Rekursionstiefe definiert wurde, damit man das Problem des unendlichen Testraumes beheben kann. Dieser Ansatz erlaubt es aber leider nicht, dass eine ideale Abdeckung (was das ist wird später genau definiert) gewährleistet werden kann. Insbesondere komplexere, zyklische Strukturen werden hierbei von dem automatischen Tool nicht getestet da die Rekursionstiefe dies oft nicht erlaubt. Mit der Nutzung

von graphspezifischen Algorithmen ist es jedoch möglich Graphabdeckungen zu ermitteln auch wenn diese eine zyklische Struktur haben und somit kann algorithmisch das Problem des Papers gelöst werden. In dieser Arbeit sollen diese graphspezifischen Algorithmen implementiert werden und dann mittels Datengeneratoren eigenständig Tests erzeugen.

Testauswertung

Die Auswertung der Tests ist im Paper darauf basierend, dass die zu testende API vor allem auch funktionale Korrektheit überprüft wird, dies bedeutet insbesondere, dass hier die HTTP-Status Codes von Anfragen ausgewertet werden sowie GraphQL eigene Statusmeldungen untersucht werden. Ein richtiger Abgleich im Test findet nicht statt, es wird lediglich überprüft ob der Rückgabe-Datentyp dem erwarteten Datentyp entspricht. Es wäre jedoch besser wenn nicht nur der Rückgabe-Datentyp stimmt sondern auch der Inhalt in diesem Datentyp. In dieser Arbeit soll das Programm mittels eines Orakels verbessert werden. Dies bedeutet, dass die GraphQL API mit Testdaten befüllt wird und Anfragen gestellt werden können, die sich dann auch logisch ineinander auflösen. Hierbei ist insbesondere zu beachten, dass Kreisstrukturen sich in gewisser Weise in sich selbst auflösen, d.h. Eingabeobjekt = Ausgabeobjekt.

0.2.2 Umsetzung

Zuallererst wird in dieser Arbeit etwas Theorie definiert und in Bezug gesetzt. Angefangen mit einer allgemeinen Definition eines Graphens im mathematischen Sinne und GraphQL als Schnittstelle, wird dann ein Bezug dieser beiden Themen zueinander hergestellt. Wenn der Bezug von Graphentheorie und GraphQL klar ist, wird der eigentliche Algorithmus für die ideale Abdeckung des Graphens algorithmisch erklärt, eine Anwendung auf GraphQL in der Theorie gezeigt und bewiesen. Danach werden die theoretischen Erkenntnisse in einem praktischen Projekt umgesetzt. Hierbei wird ein Tool erstellt, welches auf Grundlage des Überdeckungsalgorithmus einem Datengenerator Tests erstellt und dann auswertet mittels den alten bzw. erweiterten Abgleichsmethoden. Um zu zeigen, dass das neue Verfahren eine Verbesserung darstellt wird dann ein Benchmark Test zwischen altem und neuem System erstellt. Bei diesem Benchmark Test werden 3 verschiedene GraphQL-Schemas getestet, 2 Schemas aus dem alten Paper, hier ergibt sich ein direkter Vergleich an z.B. generierten Tests und Graphabdeckung. Im 3ten Schema wird ein speziell sehr zyklisches Schema ausgewertet um zu zeigen, dass einerseits die Implementierung den Algorithmus korrekt umsetzt und wie groß die Verbesserung in solchen Schemas dann sind.

Zusammenhang Graphentheorie und GraphQL

GraphQL erlaubt es uns, Typen zu definieren. Ein Type beinhaltet immer mindestens eine Property. Ein Type kann mit einem Knoten eines Graphens gleichgesetzt werden. Man nehme zum Beispiel ein Buch und definiere hierfür einen Type

```
type Book {  
  id: Int  
  title: String  
}
```

Es existiert jetzt ein Objekt Book mit den Eigenschaften id als Integer und title als String. Repräsentiert als Graphen hätten wir jetzt folgende Struktur

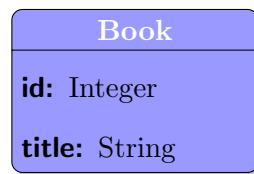
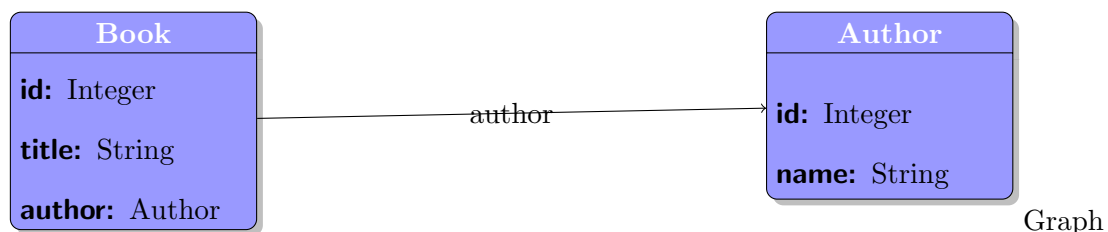


Abbildung 0.1: Graph mit 1 Type

Eine Property kann entweder ein Standardtyp sein oder auf einen Type verweisen, dies kann der eigene Type oder auch ein anderer Type sein. Fügen wir unserem Beispiel des Buches eine Property hinzu mit dem Type Author wobei der Author selbst wie folgt definiert wird:

```
type Book {  
  id: Int  
  title: String  
  author: Author  
}  
  
type Author {  
  id: Int  
  name: String  
}
```

So haben wir einen Graphen definiert, indem jedem Buch noch ein Author hinzugefügt wird. Hierdurch ergibt sich folgender Graph



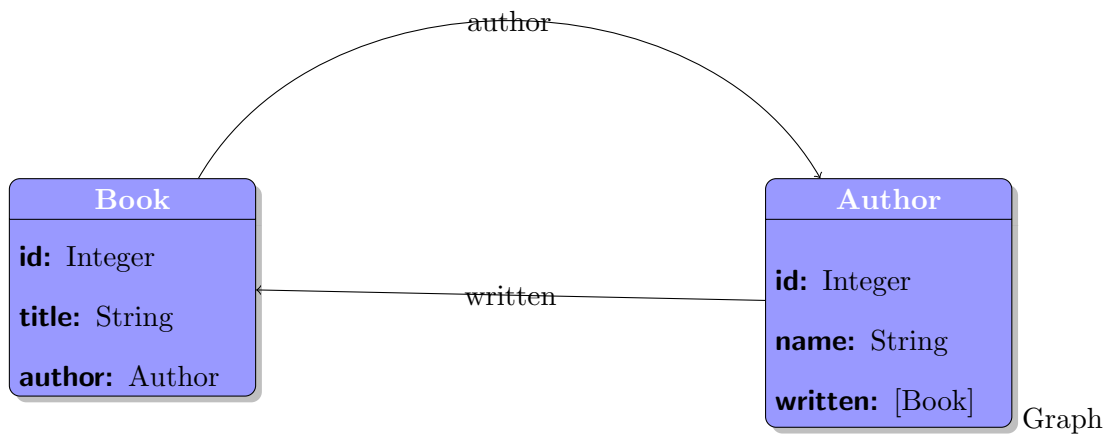
Graph

mit 2 Types

Stand jetzt, sind noch keine zirkulären Abfragen möglich. Bauen wir nun ein zirkuläres Verhalten ein indem wir dem Type Author noch die Property "written"(also alle seine geschriebenen Bücher) hinzufügen

```
type Book {  
  id: Int  
  title: String  
  author: Author  
}  
type Author {  
  id: Int  
  name: String  
  written: [Books!]  
}
```

so ergibt sich, dass wir einen zirkulären Graphen haben mit folgender Struktur



mit 2 Types

Jetzt ist es möglich, zirkuläre Abfragen in folgender Form zu stellen:

```
{  
  book{  
    author{  
      book{  
        author{  
          name  
        }  
      }  
    }  
  }  
}
```



```
    }  
  }
```

oder auch

```
{  
  book{  
    author{  
      name  
    }  
  }  
}
```

vorrausgesetzt der abzufragende Author hat nur ein Buch geschrieben. Auf Grundlage dieser hier definierten Bildungsstrukturen können dann Graphen beliebiger Größe definiert werden.

Praxis

Im folgenden wird die praktische Umsetzung der vorher erarbeiteten Theorie behandelt. Hierzu wurde/wird ein Tool geschrieben, das aufgrund eines GraphQL-Schemas automatisiert Tests erzeugt. Diese Tests werden dann ausgeführt und ausgewertet mit den entsprechenden Verbesserungen.

Toolchain

Da GraphQL ein Standard für diverse Sprachen ist und das mocken von Daten essentiell zum testen ist, kann der Teil der Testgenerierung und Auswertung nur sprachspezifisch stattfinden. Es können somit nicht alle Sprachen berücksichtigt werden. Da GraphQL vor allem in der Webentwicklung verwendet wird, bezieht sich das Testtool auf JavaScript/TypeScript mit der Testbibliothek Jest. Als Server für die Verarbeitung wird ApolloServer genutzt, es ist jedoch denkbar, dass man jeden Server einsetzen kann insofern dieser eine `executeOperation()` implementiert die einen String als Query akzeptiert. Die benötigte Toolchain ist also sehr klein, sie benötigt nur Jest für die Ausführung der generierten Tests und ApolloServer für die Ausführung der GraphQL-Resolver.

Requirements an das Tool

To Do

0.3 Glossar

Im Text werden einige Fachbegriffe genutzt. Hier findet sich deren Erklärung

Begriff Erklärung

GraphQL Waren-Managment-System; Ein System das das Lager verwaltet und die kompletten Betriebsprozesse eines Lagers abbilden kann

Onlinere Ressourcen wurden im Juli 2021 auf ihre Verfügbarkeit hin überprüft.