

Automatische Testdatengenerierung für komplexe Formulare in Geschäftsanwendungen

Das Labyrinth und der Drache der Widersprüche

Je umfangreicher und komplexer Formulare in einer Anwendung werden, desto aussichtsloser ist die manuelle Bereitstellung von Testdaten. mgm hat deshalb ein Verfahren entwickelt, das vollständig automatisiert Testdaten für die entwicklungsbegleitende Qualitätssicherung erzeugt. Ein vielversprechender Nebeneffekt: Mit dem Verfahren lassen sich auch Widersprüche in fachlichen Modellen aufspüren.

Dieses Feld darf nur angegeben werden, wenn „es sich um eine Organgesellschaft handelt und ein entsprechender Übernahmegewinn in den Einzelangaben erklärt wurde und die Anmeldung vor dem 11.11.2019 erfolgte“. So lauten die Bedingungen für ein einziges Feld des Formulars für die Körperschaftssteuer 2019. Eines Formulars wohlgermerkt, das aus 197 Seiten besteht.

Das Beispiel lässt erahnen, wie umfangreich und komplex Online-Formulare von Geschäftsanwendungen werden können. Es gibt Bedingungen und Abhängigkeiten zwischen den Feldern. Und es gibt fachliche Regeln und Zusammenhänge, die festlegen, welche Eingaben erlaubt sind und welche nicht.

All das sind Herausforderungen bei der Erzeugung von Testdaten – auch wenn sie keine fachlich sinnvollen Steuererklärungen sein müssen. Testdaten bestehen vielmehr aus Zufallswerten. Oder sie loten Maximalwerte aus und prüfen, ob die Eingabe aller erlaubten Zeichen funktioniert. Und dennoch: Wie findet man für so ein umfangreiches Formular automatisiert einen gültigen Datensatz?

Von hinten gedacht

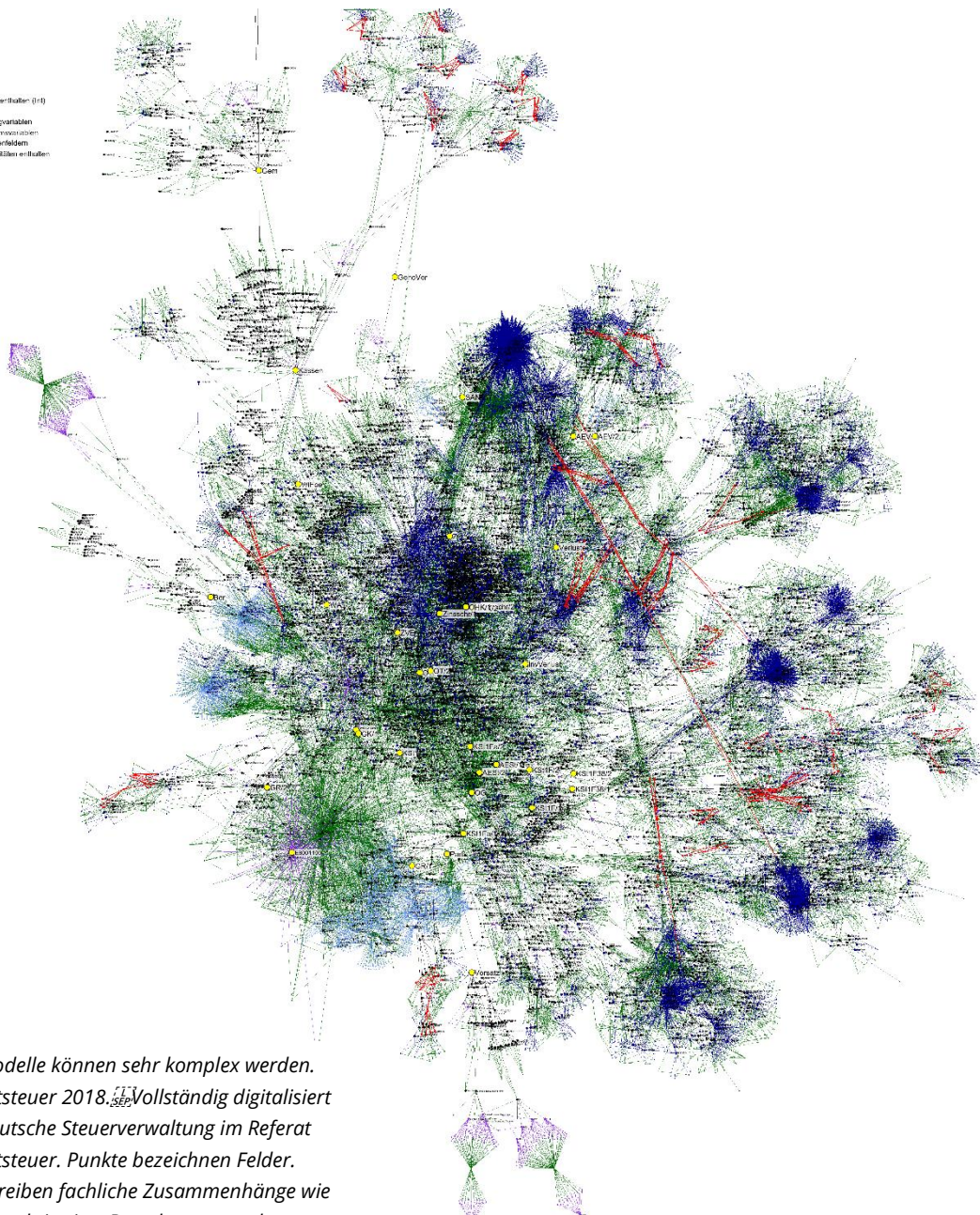
„Uns war früh klar, dass wir eine generische Lösung brauchen.“, erinnert sich Johannes Bauer. Der Diplom-Mathematiker leitet bei mgm seit einigen Jahren das Forschungsprojekt für die Testdatengenerierung.

„Wir hatten zwar schon seit Jahren einige Verfahren zur Testdatengenerierung im Einsatz. Sie kamen allerdings nicht komplett ohne Fachwissen aus. Das war immer wieder arbeits- und zeitaufwendig.“

Zufällig Testdaten zu erzeugen, war keine Option. „Für ein sehr umfangreiches Formular in einem Kundenprojekt haben wir ausgerechnet, wie hoch die Wahrscheinlichkeit liegt, zufällig einen Datensatz zu finden, der eine gültige Kombination der angegebenen und nicht angegebenen Felder enthält. Das Ergebnis war 10^{-300} Prozent. Die Wahrscheinlichkeit ist astronomisch gering“, sagt Bauer.

KSt E30 2019 6.1

- Vordrucke und Schlüsselfelder
- Rückgangswahlwert (ROCI)
- String: Indikationskriterien (Ind)
- Datum: Jahr/Monat/Tag (jmt)
- Zahlendollar (lit)
- Zahlendollar: in Nicht-Beständen erlauben (i-r)
- Constante
- Integervergleiche zwischen Stringvariablen
- Integervergleiche zwischen Datenassoziationen
- Integervergleiche zwischen Zahlendollern
- Integervergleiche, die Nicht-Beständen erlauben



Fachliche Modelle können sehr komplex werden. Körperschaftsteuer 2018, ^{1.1.2018} ~~2018~~ vollständig digitalisiert durch die deutsche Steuerverwaltung im Referat Körperschaftsteuer. Punkte bezeichnen Felder. Linien beschreiben fachliche Zusammenhänge wie z. B. Ausschlusskriterien, Berechnungen oder Wertevergleiche.

Anstatt bei der Testdatenerzeugung „von vorne“ vorzugehen – also erst Daten zu generieren und dann zu schauen, ob sie passen – setzt der Ansatz „von hinten“ an. Was muss erfüllt sein, damit ein Feld überhaupt angegeben sein darf? Die Antwort darauf liegt genau in den fachlichen A12-Modellen des modellbasierten Ansatzes. Sie repräsentieren die hinter den Feldern eines Formulars liegenden Datenvorgaben sowie Validierungsregeln und Berechnungen. Sprich: die Fachlichkeit der Anwendung.

Entscheidungsproblem: die richtigen Abzweigungen finden

Auch wenn die fachlichen Modelle gegeben sind, bleiben eine Reihe von Herausforderungen für die Generierung gültiger Testdatensätze. Viele Bedingungen, die für einzelne Felder definiert sind, eröffnen mehrere Möglichkeiten, gültige Testdaten zu erzeugen. Man kann sich diese Möglichkeiten wie Abzweigungen eines Weges vorstellen. Für jede Abzweigung muss eine Entscheidung getroffen werden, welcher Weg beschritten werden soll. Die meisten Kombinationen aus Abzweigungen erzeugen jedoch irgendwann Widersprüche, sie münden in Sackgassen. Es ist so, als müsste man einen der wenigen Wege aus einem gigantischen Labyrinth herausfinden.

In einem Kundenprojekt ergaben sich aus einem sehr umfangreichen fachlichen Modell alleine 212.000 Möglichkeiten. Es ist nicht annähernd möglich, sie alle durchzuprobieren. Wir würden eine halbe Ewigkeit durch das Labyrinth laufen.

Auswege aus dem Labyrinth: SMT-Solver

Zum Glück gibt es spezielle Algorithmen, die genau auf solche Problemstellungen spezialisiert sind.

Sie gehen den Fragen nach,

- in welcher Reihenfolge man die Entscheidungen treffen sollte,
- wie man schnell herausfindet, welche Abzweigung eine Sackgasse ist
- und wie man bei Zusatzbedingungen frühere Kenntnisse verwerten kann und zu welchen Abzweigungen man zurückgehen muss.

Diese Algorithmen gehören in der Informatik und Mathematik zum Bereich der sogenannten „satisfiability modulo theories“ (SMT). Sie beschäftigen sich mit Entscheidungsproblemen, bei denen es um die Lösbarkeit von logischen Formeln geht. Für den Einsatz in der Praxis gibt es spezielle SMT-Solver, die die Algorithmen in Bibliotheken bereitstellen. Und es gibt [eine standardisierte Sprache – SMT-LIB](#) – für die Form des Inputs der Modelle und des Outputs der Daten.

Um für die Testdatenerzeugung von Anwendungen, die mit der A12 Plattform entwickelt wurden, einen SMT-Solver zu nutzen, war zunächst eine Übersetzung der fachlichen Regeln nötig. Dafür wurde im Rahmen des Forschungsprojekts eine Brücke von der A12 Regelsprache zu SMT LIB entwickelt. Die Regeln für ein Formular werden damit zunächst nach SMT-LIB übersetzt. Die Übersetzung wird als Input durch den SMT-Solver

gejagt. Der Output des SMT-Solvers wird anschließend wieder zurückübersetzt in einen Datensatz im Format der A12 Regelsprache.

Hundertprozentige Abdeckung der Felder

Die Testdatenerzeugung ist ein sehr rechenintensiver Vorgang. Er gleicht der Suche nach einem schmalen befestigten Weg durch ein Sumpfgebiet. Der kleinste Fehltritt kann in unendlicher Laufzeit münden. Auch wenn die SMT-Solver immer weiter verbessert werden, kann die Testdatenerzeugung bei sehr großen Modellen mit komplexen Regelwerken Stunden in Anspruch nehmen. Die optimale Übersetzung von A12-Regeln nach SMT-LIB sowie die beste Einstellung und Verwendung der SMT-Solvers ist das Ergebnis jahrelanger Forschung und kontinuierlicher Verbesserung.

„Wir müssen regelmäßig in die Verbesserung der Performance investieren. Die Modelle von Geschäftsanwendungen werden typischerweise sukzessive umfangreicher und komplexer“, sagt Bauer. „Heute benötigen wir nur noch 20 Prozent der Zeit für die Erzeugung der Testdaten im Vergleich zu früheren Lösungen. Dabei sind 100 Prozent der Felder abgedeckt. Der entscheidende Vorteil des neuen Verfahrens ist aber, dass es vollständig automatisiert abläuft. Für die Generierung ist kein Fachwissen über die Inhalte der Formulare mehr erforderlich. Wir nutzen direkt das Wissen aus den fachlichen Modellen.“

Eine Lösung für das „Multihopp-Problem“

„Bei der Entwicklung des Generators für die Testdatenerzeugung haben wir gemerkt, dass das Verfahren ein weiteres Problem lösen kann.“, sagt Bauer. „Es erkennt Widersprüche in den Regeln, die dafür sorgen, dass manche Felder nicht angegeben werden können.“

Ein einfaches Beispiel: Regel 1 besagt, dass die Felder A und B eines Formulars gemeinsam angegeben sein müssen. Regel 2 besagt, dass die Felder A und B nicht gemeinsam angegeben sein dürfen. Als Folge können Anwender für die betroffenen Felder keine Eingaben machen. Sie sind durch den Regelwiderspruch unbrauchbar geworden.

In der Praxis fallen solche einfachen Widersprüche direkt auf. Es gibt jedoch deutlich schwierigere Fälle, die nur sehr schwierig aufzufinden sind. „Wir waren selber erstaunt, wie versteckt manche Widersprüche sein können, die sich über mehrere Regeln erstrecken. In der Praxis hat die Testdatenerzeugung zum Beispiel einen Widerspruch aufgedeckt, der sich in einem Steuerformular über einen Zusammenhang von acht Regeln und acht Feldern erstreckt hat.“, sagt Bauer. „Intern bezeichnen wir dieses Phänomen als Multi-Hopp-Problem. Ein Mensch hat kaum eine Chance, solche Fehler zu entdecken.“

Widersprüche direkt bei der Modellierung erkennen

Das entwickelte Verfahren erkennt nicht nur, dass solche Widersprüche existieren. Es legt auch offen, welche Regeln und Felder betroffen sind. „Für eine fehlerfreie Modellierung sind diese Informationen von unschätzbarem Wert.“, sagt Bauer. „Wir wollen sie deshalb auch möglichst früh im Entwicklungsprozess den Business Analysten und Fachexperten, die die Regeln erstellen, zur Verfügung stellen.“

Das Verfahren wird zurzeit im Rahmen der Forschung und Entwicklung der Enterprise Low Code Plattform A12 in die Modellierungswerkzeuge integriert. Künftig können die Werkzeuge auf solche Widersprüche direkt bei der fachlichen Modellierung hinweisen und sicherstellen, dass die Modelle mathematisch gesehen in Ordnung sind.

„Es ist ein schönes Beispiel, wie die Ergebnisse eines internen Forschungsprojekts zweitverwertet werden können.“, resümiert Johannes Bauer. „Wir haben gewissermaßen nicht nur die Wege aus dem Labyrinth gefunden. Sondern im Vorbeigehen auch gleich den darin hausenden Drachen gezähmt.“