



FACHGEBIET KOMPLEXE UND VERTEILTE IT-SYSTEME



**Maschinelle Extraktion von Modes aus Anforderungsdokumenten  
zur Unterstützung von Entwicklern bei der Modellierung von  
globalen Zustandsmodellen**

**- EXPOSÉ -**

Name, Vorname	Pudlitz, Florian
Erstellung des Exposés	25. August 2016
Studiengang	Computer Engineering (M.Sc.)
Betreuer	Prof. Dr. Vogelsang, Jonas Winkler

# 1 Einleitung

Die stetig steigende Anzahl von Funktionen und Assistenzsystemen in einem Fahrzeug führt zu einem Komfortgewinn für den Fahrer und steigert die Sicherheit im Straßenverkehr. Dies geht einher mit einem immer wachsenden Softwareanteil, der für die Steuerung, Regelung, Überwachung und Automatisierung im Fahrzeug verantwortlich ist [1]. Die wachsende Komplexität stellt eine zunehmende Herausforderung für die Entwickler und Tester dar. Mit detaillierten Beschreibungen der Einzelfunktionen in den Anforderungsdokumenten muss ein funktionierendes Gesamtsystem entwickelt werden, in dem alle Komponenten zuverlässig miteinander arbeiten.

Der Vereinfachung und Übersichtlichkeit dieser komplexen Entwicklungssysteme dient die modellbasierte Softwareentwicklung. Hierbei wird mit Hilfe von Modellen die Software modelliert und anschließend automatisiert in Quellcode umgesetzt. Die Modelle werden aus Anforderungsdokumenten abgeleitet, indem verschiedene Zustände extrahiert, strukturiert und in Beziehung zueinander gesetzt werden. Üblicherweise sind Anforderungsdokumente, beispielsweise Lastenhefte, in natürlicher Sprache mit domänenspezifischen Begriffen und Abkürzungen verfasst. Das systemgestützte, automatisierte Suchen von Zuständen findet aufgrund der Komplexität der verwendeten Sprache zum aktuellen Zeitpunkt kaum Anwendung. Ursächlich hierfür ist die unzureichende Betrachtung und Entwicklung von speziellen Extrahierungsverfahren.

Ein funktionierendes automatisiertes Verfahren dient der Unterstützung der Entwickler, den Gesamtkontext eines umfangreich beschriebenen Systems zu erfassen und Modelle schneller strukturieren und besser modellieren zu können.

## 2 Modes

In der modellbasierten Softwareentwicklung ist die Abbildung der Software und deren Funktionen auf unterschiedliche Systemzustände enthalten. Diese Zustände werden Modes genannt [4]. Sie sind komponentenübergreifend und bildet den Zustand des Gesamtsystems ab.

Die folgenden Kriterien müssen erfüllt sein, um als Mode in einem Lastenheft erkannt zu werden:

- Der Wechsel von Modenzuständen ist nur während des Betriebs möglich. Parameter und Konstanten werden nicht als Mode erkannt.

- Die Zustände können ausschließlich diskrete Werte annehmen. Kontinuierliche Zustandsübergänge, beispielsweise Temperatur und Geschwindigkeit, sind ausgeschlossen.
- Die Zustandsänderung ist von außen durch den Nutzer beobachtbar. Es werden keine internen Änderungen berücksichtigt.
- Modes müssen eine funktionale Bedeutung für das Gesamtsystem haben. Fehlerzustände sind keine Modes.

Modes umfassen mehrere einzelne Komponenten und werden daher auch globale Modes genannt.

### 3 Aufgabenstellung

In dieser Arbeit soll ein Verfahren zur Erkennung globaler Modes entwickelt werden, das anschließend mit Hilfe eines Tools evaluiert wird.

Folgende Teilaufgaben wurden identifiziert:

- Entwicklung eines Verfahrens zur Erkennung von globalen Modes.
- Identifikation der möglichen Zustände, die ein Mode annehmen kann.
- Entwicklung eines Softwaretools zur Umsetzung des Verfahrens.
- Evaluation der Ergebnisse durch einen Vergleich der Programmausgaben und einer manuellen Analyse.

### 4 Lösungsansatz

Um Modes in Lastenheften automatisiert zu erkennen, sind umfangreiche Analysen der Anforderungs- und Informationstexte nötig. Mit Hilfe von Techniken aus dem Bereich der natürlichen Sprachverarbeitung sollen Modes erkannt werden [3][2].

Des Weiteren werden mehrere unterschiedliche Lastenhefte analysiert um zu erkennen, welche Modes global sind. Eine besondere Schwierigkeit stellen Modes dar, deren Bedeutung gleich ist. Sie müssen durch Häufigkeitsanalysen zusammengefasst werden. Da die Ergebnisse nicht zu 100% korrekt sind, soll ebenfalls der Benutzer in den Mode-Auswahlprozess einbezogen werden. Die möglichen Zustände der Modes werden über die grammatikalischen Zusammenhänge erkannt.

Um die Qualität weiter zu erhöhen, wird das Ergebnis für weitere Analysen berücksichtigt. So kann die Zuverlässigkeit der Modeerkennung mit der Zeit stetig verbessert werden. Weitere Optimierungen, beispielsweise verminderte Ausführungszeit, sind denkbar.

## 5 Evaluation

Um das Verfahren anschließend zu evaluieren, erfolgt die Implementierung in einem Softwaretool und der Vergleich der Ausgaben mit einer Referenzliste. Dazu wird ein Set von Lastenheften ausgewählt, in denen globale Modes enthalten sind. Die Lastenhefte werden von der Daimler AG zur Verfügung gestellt. Die Referenzliste von Modes entsteht durch eine manuelle Analyse der Lastenhefte. Dafür wird eine Befragung einer kleinen Gruppe von Softwareentwicklern und Testern der Daimler AG durchgeführt. Anschließend können die identifizierten Modes mit den Ergebnissen des Softwaretools verglichen werden. Die Qualität der Evaluation kann verbessert werden, wenn die Befragungen mit weiteren Sets von Lastenheften durchgeführt wird.

## Literatur

- [1] Alexander Pretschner, Manfred Broy, Ingolf H. Kruger, and Thomas Stauner. Software engineering for automotive systems: A roadmap. In *2007 Future of Software Engineering*, FOSE '07, pages 55–71, Washington, DC, USA, 2007. IEEE Computer Society.
- [2] Helmut Schmid. Trace prediction and recovery with unlexicalized pcfgs and slash features. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics*, ACL-44, pages 177–184, Stroudsburg, PA, USA, 2006. Association for Computational Linguistics.
- [3] Kristina Toutanova and Christopher D. Manning. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the 2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora: Held in Conjunction with the 38th Annual Meeting of the Association for Computational Linguistics - Volume 13*, EMNLP '00, pages 63–70, Stroudsburg, PA, USA, 2000. Association for Computational Linguistics.

- [4] A. Vogelsang, H. Femmer, and C. Winkler. Systematic elicitation of mode models for multifunctional systems. In *2015 IEEE 23rd International Requirements Engineering Conference (RE)*, pages 305–314, Aug 2015.