
Entwicklung von Testauswahlmethoden für die Absicherung von Änderungen auf Gesamtfahrzeugebene

Vom Fachbereich Maschinenbau an der
Technischen Universität Darmstadt
zur Erlangung des Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

Dissertation

vorgelegt von

Christina Theresia Singer, M. Sc.
aus Baiersdorf

Berichterstatter: Prof. Dr. rer. nat. Hermann Winner

Mitberichterstatter: Prof. Dr.-Ing. Joachim Metternich

Tag der Einreichung: 28.09.2015

Tag der mündlichen Prüfung: 16.12.2015

Darmstadt 2015

D 17

Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftliche Mitarbeiterin des Fachgebiets Fahrzeugtechnik (FZD) der Technischen Universität Darmstadt.

Ich danke herzlich Herrn Prof. Dr. rer. nat. Hermann Winner, dem Leiter des Fachgebiets Fahrzeugtechnik, der mit kritischen Diskussionen, großen Freiräumen und Vertrauen in meine Fähigkeiten zum Gelingen der Arbeit und zu meiner persönlichen Entwicklung beigetragen hat.

Herrn Prof. Dr.-Ing. Joachim Metternich, stellvertretender Institutsleiter des Instituts für Produktionsmanagement, Technologie und Werkzeugmaschinen, danke ich für die freundliche Übernahme des Korreferats.

Darüber hinaus bedanke ich mich bei allen Mitarbeitern von FZD, einschließlich der Werkstätten und des Sekretariats, für die angenehme Zusammenarbeit sowie das freundliche Arbeitsumfeld.

Ganz besonders danken möchte ich meinem Ehemann Matthias sowie meinen Kindern Anna-Sophie, Leonard und Ferdinand. Ihre Unterstützung und ihr Verständnis waren der Grundstein für meine Motivation und das Gelingen dieser Arbeit.

Christina Singer

Baiersdorf, Februar 2016



Inhaltsverzeichnis

Abkürzungen	VII
Formelzeichen und Indizes	VIII
Kurzfassung	X
1 Einleitung und Motivation	1
2 Gesamtfahrzeugabsicherung in der Automobilindustrie	3
2.1 Begriffsdefinition.....	3
2.2 Einordnung in den Entwicklungsprozess	4
2.3 Anforderungen an die Absicherung.....	6
2.4 Durchführung der Absicherung	7
2.5 Paradigmen der Gesamtfahrzeugabsicherung	9
2.5.1 Paradigma 1: Spezifikationsbasierte Absicherung	9
2.5.2 Paradigma 2: Feldbasierte Absicherung.....	10
2.5.3 Vergleich der Paradigmen	10
2.6 Herausforderungen der Gesamtfahrzeugabsicherung.....	13
3 Absicherung von Änderungen in der Automobilindustrie	15
3.1 Änderungen	15
3.2 Änderungsmanagement	17
3.3 Anforderungen an das Änderungsmanagement.....	19
3.4 Testauswahlmethoden zur Absicherung von Änderungen.....	20
3.4.1 Manuelle Testauswahl.....	20
3.4.2 Automatisierte Testauswahl	22
3.5 Herausforderungen bei der Absicherung von Änderungen	30
3.6 Präzisierung der Ausgangslage, der Aufgabenstellung und des Vorgehens.....	31
4 Prinzipielle Lösungsansätze für Testauswahlmethoden	33
4.1 Definition Testauswahlmethode	33
4.2 Abstraktion der Testauswahlsituation.....	33
4.3 Prinzipielle Testauswahlmethoden	35
4.3.1 Inklusionsbasierte Testauswahl	36
4.3.2 Exklusionsbasierte Testauswahl.....	36
4.4 Bewertung prinzipieller Testauswahlmethoden.....	37
4.4.1 Selektionsqualität	38
4.4.2 Selektionsergebniskosten	40
4.4.3 Selektionsaufwand	43

4.4.4 Umgang mit mehreren Änderungen.....	45
4.5 Zusammenfassung und Fazit.....	50
5 Inklusionsanalyse	53
5.1 Prinzip der Inklusionsbeweissführung.....	53
5.2 Stand der Technik.....	55
5.3 Fazit.....	58
6 Exklusionsanalyse	60
6.1 Prinzip der Exklusionsbeweissführung.....	60
6.2 Exemplarische Umsetzung	63
6.2.1 Ausschlussbeweis 1 für Softwareänderungen.....	63
6.2.2 Ausschlussbeweis 1 für Hardwareänderungen	65
6.2.3 Ausschlussbeweis 2 für Software- und Hardwareänderungen.....	68
6.3 Fazit.....	70
7 Entwicklung eines Gesamtkonzeptes für die Testauswahl.....	72
7.1 Vergleich von Inklusion und Exklusion	72
7.2 Gesamtkonzept für die Testauswahl.....	74
7.3 Ergebniserwartung.....	76
8 Validierungsansatz für das Testauswahlkonzept	80
8.1 Untersuchungsziele	80
8.2 Evaluierung der Güte der Testauswahl.....	81
8.3 Evaluierung der Effizienz der Testauswahl.....	84
8.4 Evaluierung der Generalität der Testauswahl.....	86
8.5 Fazit.....	87
9 Schlussfolgerungen für den Entwicklungsprozess.....	89
9.1 Design for Exclusion.....	89
9.2 Anforderungen an das Testdesign	91
9.3 Fazit und Diskussion	93
10 Zusammenfassung und Ausblick.....	96
Literaturverzeichnis	102
Eigene Veröffentlichungen	111
Betreute studentische Arbeiten.....	112

Abkürzungen

Abkürzung	Beschreibung
<i>AB</i>	Ausschlussbeweis
<i>ASIL</i>	Automotive Safety Integrity Level
<i>ATZ</i>	Automobiltechnische Zeitschrift
<i>CAD</i>	Computer Aided Design
<i>CAE</i>	Computer Aided Engineering
<i>CDM</i>	Component-Dependency-Model
<i>C-FAR</i>	Change Favorable Representation
<i>CIA</i>	Change Impact Analysis
<i>CPM</i>	Change Propagation Method
<i>CR</i>	Change Request
<i>DSM</i>	Design Structure Matrix
<i>EB</i>	Einschlussbeweis
<i>EMV</i>	Elektromagnetische Verträglichkeit
<i>ESC</i>	Electronic Stability Control
<i>FEM</i>	Finite-Elemente-Methode
<i>FMEA</i>	Failure Modes and Effects Analysis
<i>FO</i>	Funktionsorientierung
<i>HiL</i>	Hardware-in-the-Loop
<i>OEM</i>	Original Equipment Manufacturer
<i>QFD</i>	Quality Function Deployment
<i>ROC</i>	Receiver Operating Characteristic
<i>RTS</i>	Regression Test Selection
<i>SiL</i>	Software-in-the-Loop
<i>SOP</i>	Start of Production
<i>SPICE</i>	Software Process Improvement and Capability Determination
<i>SQ</i>	Selektionsqualität
<i>UML</i>	Unified Modeling Language

Formelzeichen und Indizes

Symbol	Einheit	Beschreibung
A	-	Änderung
b	-	Auswirkungsgröße der Änderung
c	-	Auswirkungsbegrenztheit der Änderung
C	-	Gewichtungsfaktor für Falschklassifizierungskosten
d	-	Physikalischer Parameter
f	-	Anzahl falscher Tests
F	-	Funktion
k	-	Anzahl der Wechselwirkungen
K	-	Kostenfaktor
n	-	Anzahl der Änderungen
N	-	Menge der insgesamt notwendigen Tests zur Absicherung eines Systems
N'	-	Menge der notwendigen Tests zur Absicherung der Änderung
$NSEK$	-	Normalisierte Selektionsergebniskosten
P	-	Wahrscheinlichkeit
r	-	Anzahl richtiger Tests
R	-	Ergebnis („result“)
RN	-	Menge der richtig negativen Tests
RP	-	Menge der richtig positiven Tests
SA	-	Selektionsaufwand
SEK	-	Selektionsergebniskosten
SQ	-	Selektionsqualität
t	-	Testfall
T	-	Menge der durchgeführten Tests zur Absicherung eines Systems
T'	-	Menge der durchgeführten Tests zur Absicherung der Änderung
v	m/s	Geschwindigkeit
z	-	Testziel
μ	-	Reibwert

Index	Beschreibung
<i>A</i>	Auswahl
<i>D</i>	Durchführung
<i>DG</i>	Dunkelgrau
<i>ex</i>	Exklusion
<i>ges</i>	Gesamt
<i>f</i>	Falsch
<i>F</i>	Fehler
<i>HG</i>	Hellgrau
<i>i</i>	Index
<i>j</i>	Index
<i>in</i>	Inklusion
<i>inital</i>	initial (anfänglich)
<i>l</i>	Index
<i>m</i>	Index
<i>max</i>	Maximum, maximal
<i>n</i>	Negativ
<i>p</i>	Positiv
<i>r</i>	Richtig
<i>red</i>	Reduziert
<i>S</i>	Schwarz
<i>simp</i>	Simplifiziert
<i>W</i>	Weiß
<i>w</i>	Wahr

Kurzfassung

Die wachsende Anzahl von Softwarefunktionen sowie die erhöhte Varianten- und Ausstattungsvielfalt heutiger Fahrzeuge führen zu einer gesteigerten Komplexität automobiler Systeme. Begleitet wird dies von stetig wachsenden Anforderungen an die Qualität und Zuverlässigkeit der Fahrzeuge. Zusammengenommen folgt daraus ein enormer Anstieg des Aufwands für die Gesamtfahrzeugabsicherung. Der hohe Innovationsdruck in der Automobilindustrie führt zudem dazu, dass der effiziente Umgang mit kurzfristigen Änderungen ein entscheidender Wettbewerbsfaktor geworden ist. Um sicherzustellen, dass die Qualität und Sicherheit des Systems durch die Modifikation nicht beeinträchtigt wird, muss der Absicherungsprozess revidiert werden. Die Ermittlung des benötigten Wiederholungstestbedarfs ist dabei die zentrale Herausforderung. Gegenstand dieser Arbeit ist daher die Untersuchung der Fragestellung, ob eine Testauswahlmethode existiert, die eine Identifikation der erforderlichen Tests im Fall von Änderungen in der Gesamtfahrzeugabsicherung ermöglicht.

Dazu werden im ersten Teil der Arbeit die Rahmenbedingungen für die zu entwickelnde Testauswahlmethode analysiert. Hierzu zählen die Gegebenheiten der Gesamtfahrzeugabsicherung und des automobilen Änderungsmanagements sowie der Stand der Technik.

Die Entwicklung der Testauswahlmethode erfolgt darauf aufbauend anhand einer deduktiven Vorgehensweise. Dabei wird die Fragestellung zunächst abstrahiert. Anhand des generalisierten Problems werden dann zwei prinzipielle Lösungsmöglichkeiten aufgestellt. Der Inklusionsansatz identifiziert aus der Menge aller Tests die notwendigen Testfälle, wohingegen der Exklusionsansatz die nicht notwendigen Testfälle ermittelt. Ein grundsätzlicher Vergleich anhand der Kriterien Selektionsqualität, Selektionskosten, Selektionsaufwand sowie Umgang mit mehreren Änderungen zeigt, dass unter der Annahme einer idealen Selektionsqualität auf dieser Abstraktionsebene beide Ansätze als gleichwertig einzustufen sind.

Im nächsten Schritt erfolgt daher eine weitere Konkretisierung. Diese legt dar, dass in der Literatur viele Konzepte existieren, die eine effiziente Auswahl von Testfällen mit Hilfe der Inklusion ermöglichen. Großer Nachteil ist jedoch, dass die Güte der Ergebnisse von der Qualität des zugrundeliegenden Modells zur Ermittlung der Wechselwirkungen einer Änderung abhängig ist. Für die Exklusion existieren in der Literatur keine beispielhaften Konkretisierungen, daher wird eine exemplarische Umsetzung dieses Konzeptes anhand von zwei Ausschlussbeweisen entwickelt. Die entworfene Realisierung gewährleistet eine hohe Vollständigkeit der Testauswahl, so dass die resultierende

Testmenge immer sicher (i.S.v. keine falsch negative Testauswahl) ist. Dies macht eine Anwendung im rechtlich relevanten Bereich der Gesamtfahrzeugabsicherung möglich. Nachteilig ist jedoch der hohe Aufwand der Ausschlussbeweissführung.

Anhand einer Effizienzbetrachtung wird nachfolgend ein Gesamtkonzept für die Testauswahl entwickelt, das die spezifischen Vorteile beider prinzipiellen Lösungsmöglichkeiten vereint. Die Inklusionsanalyse wird dabei zur Vorselektion von Testfällen für die Exklusion eingesetzt. Das vorgeschlagene Gesamtkonzept ermöglicht eine sichere und effiziente Identifizierung von Testfällen, die bei der Überprüfung von Änderungen in der Gesamtfahrzeugabsicherung entfallen können.

Des Weiteren erfolgt die Entwicklung eines Validierungsansatzes. Die Güte der Testauswahl kann mithilfe von Langzeitdaten erfasst werden, die eine statistische Aussage über die Sicherheit der Ergebnisse liefern. Die Effizienz der Testauswahl kann auf Basis eines vereinfachten Kostenmodells, das aus drei Kostenfaktoren besteht, beurteilt werden. Zur Validierung der Generalität wird vorgeschlagen, die Methode anhand von exemplarischen Änderungsbeispielen auf acht unterschiedliche Testkategorien sowie zwei verschiedene Änderungsarten anzuwenden. Somit liegt ein umfassendes Validierungskonzept für die Testauswahlmethode vor, für dessen Anwendung allerdings noch Voraussetzungen fehlen.

Eine Darlegung der Schlussfolgerungen für den Entwicklungsprozess bildet den letzten Teil der Arbeit. Sie identifiziert notwendige Anforderungen der entwickelten Methode an die Softwareentwicklung sowie das Testdesign. Außerdem erfolgt eine Diskussion möglicher Alternativen.

Die Ergebnisse dieser Arbeit bilden die Grundlage für die Realisierung einer Testauswahlmethode zur Identifikation der erforderlichen Tests im Fall von Änderungen in der Gesamtfahrzeugabsicherung. Sie liefern darüber hinaus einen Beitrag zur Untersuchung der Potenziale und Defizite der Gesamtfahrzeugabsicherung im Entwicklungsprozess in der Automobilindustrie.

1 Einleitung und Motivation

Die Anzahl von Softwarefunktionen im Automobil hat in den letzten Jahrzehnten stark zugenommen. Moderne Fahrzeuge sind mit bis zu über achtzig Steuergeräten ausgerüstet, die über mehrere Bussysteme miteinander verbunden sind und gemeinsam mehrere tausend Funktionen realisieren.¹ Begründet werden kann dieser Anstieg mit einer vergrößerten Nachfrage nach neuen Funktionen (z.B. Fahrerassistenzsysteme, Fahrdynamikregelsysteme, Infotainment) sowie einer verbesserten Verfügbarkeit leistungsfähiger und kostengünstiger Hardware.² Ein weiterer Trend ist die erhöhte Varianten- und Ausstattungsvielfalt, die aus dem Wunsch der Kunden nach stärkerer Individualisierung resultiert.² Diese Entwicklungen führen zu einer gesteigerten Komplexität automobiler Systeme. Begleitet wird dies von stetig wachsenden Anforderungen an die Qualität und Zuverlässigkeit der Fahrzeuge.² Zusammengenommen folgt daraus ein enormer Anstieg des Aufwands für die Gesamtfahrzeugabsicherung, die der gesteigerten Komplexität und Variantenvielfalt sowie den höheren Anforderungen an Qualität und Sicherheit Rechnung tragen muss.

Der hohe Innovationsdruck in der Automobilindustrie erfordert zudem immer kürzere Entwicklungszeiten, so dass die für die Gesamtfahrzeugabsicherung zur Verfügung stehende Zeit immer kürzer wird. Eine weitere Folge des hohen Innovationsdruckes ist, dass sich Plattformen und Systeminfrastrukturen schnell verändern.² Die schnelle Integration neuer Technologien sowie der effiziente Umgang mit Änderungen sind daher entscheidende Wettbewerbsfaktoren in der Automobilindustrie.

Innovationen finden zu 90 % in der Elektronik und Software eines Fahrzeugs statt.³ Daher nimmt auch die Änderungshäufigkeit zu.³ Änderungen treten aufgrund von funktionalen Erweiterungen, funktionalen Modifikationen, Kostenreduktionsmaßnahmen oder Adaptionen an neue Hardware auf.⁴ Um die Auswirkungen einer Änderung zu bestimmen und um sicherzustellen, dass die Qualität und Sicherheit des Systems durch die Modifikation nicht beeinträchtigt wird, muss der Absicherungsprozess revidiert werden. Die Identifikation der Auswirkungen der Änderung sowie die Ermittlung des benötigten Wiederholungstestbedarfs sind die zentralen Herausforderungen dieser Situation. Da es sich bei dem Gesamtfahrzeug um ein hochgradig vernetztes, eingebet-

¹ Broy et al.: Architekturen softwarebasierter Funktionen im Fahrzeug, 2011

² Broy: Challenges in Automotive Software Engineering, 2006

³ Broy et al.: Architekturen softwarebasierter Funktionen im Fahrzeug, 2011

⁴ Gustavsson: Architecting Automotive Product Lines, 2010

tetes System handelt, das aus unterschiedlichsten Komponenten (Mechanik, Hydraulik, Pneumatik, Elektrik, Software etc.) besteht, erfordert die Ermittlung der Auswirkungen von Änderungen ein Verständnis des vollständigen Wirkzusammenhanges all dieser Einzelkomponenten. Die Vorhersage der Auswirkungen von Änderungen wird dabei vor allem aufgrund der aktuellen, langjährig gewachsenen Systemstruktur sowie der stetig steigenden Vernetzung erschwert.⁵

Die Handhabung von Änderungen in automobilen Absicherungsprozessen auf Gesamtfahrzeugebene ist damit eine zentrale Herausforderung heutiger Entwicklungsprojekte. Unterstützung in Form einer belastbaren, objektiv nachvollziehbaren Testauswahlmethodik sollte insbesondere auch vor dem Hintergrund von Produkthaftungsfragen vorliegen. Die dafür notwendigen Schritte und die damit verbundenen Herausforderungen werden in dieser Arbeit analysiert und Lösungsansätze entwickelt.

⁵ Broy et al.: Architekturen softwarebasierter Funktionen im Fahrzeug, 2011

2 Gesamtfahrzeugabsicherung in der Automobilindustrie

In diesem Kapitel wird in Anlehnung an Singer et al.⁶ ein Überblick über die Gesamtfahrzeugabsicherung in der Automobilindustrie gegeben. Es erfolgt zunächst eine Begriffsklärung sowie eine Einführung in die Gesamtfahrzeugabsicherung. Nach der Betrachtung der Anforderungen werden die Grundlagen der Gesamtfahrzeugabsicherung erläutert sowie unterschiedliche Paradigmen dieses Entwicklungsschrittes analysiert. Am Ende des Kapitels werden die Herausforderungen der Gesamtfahrzeugabsicherung in der Automobilindustrie diskutiert.

2.1 Begriffsdefinition

Der Begriff Gesamtfahrzeugabsicherung bezeichnet die letzte Teststufe im Entwicklungsprozess. Sie beinhaltet alle Aktivitäten, die innerhalb der Entwicklungsorganisation durchgeführt werden, um die Qualität des Gesamtfahrzeugs abschließend, d.h. vor dem Inverkehrbringen, zu beurteilen. Die Qualität charakterisiert dabei die Gesamtheit von Eigenschaften und Merkmalen des Fahrzeugs, die sich auf dessen Eignung zur Erfüllung der Erfordernisse des Kundenbetriebs beziehen.⁷ Die Absicherung des Gesamtfahrzeugs dient somit als Grundlage der Entscheidung über die Freigabe des Fahrzeugs für die Produktion und den Kundenbetrieb. Sie ist von dieser aber strikt zu trennen, da es sich bei der Freigabe um einen eigenständigen Prozessschritt handelt. Die Freigabe ist eine binäre Entscheidung (Freigabe ja oder nein), die auf den Ergebnissen der Absicherung aufbaut. Dabei ist dem Entscheider jedoch ein gewisser Ermessensspielraum gegeben, der neben den Absicherungsresultaten auch andere Randbedingungen (z.B. Projektdruck) berücksichtigt. Oftmals wird der Begriff Freigabe als Synonym für den Absicherungsprozess verwendet, weil am Ende eines erfolgreichen Produktentwicklungsprozesses immer eine erteilte Freigabe steht und somit eine Freigabe immer implizit einen Absicherungsprozess voraussetzt.

⁶ Singer et al.: Gesamtfahrzeugabsicherung, 2015

⁷ Nörenberg: Effizienter Regressionstest von E/E-Systemen, 2012, S. 22

2.2 Einordnung in den Entwicklungsprozess

Abbildung 2-1 stellt vereinfacht die Rolle der Gesamtfahrzeugabsicherung in der Produktentwicklung dar. Zu Beginn des Prozesses werden die Anforderungen aufgestellt. In einem Lastenheft legt der OEM (Original Equipment Manufacturer) u.a. fest, welche Merkmale das Fahrzeug besitzt, welche Funktionen implementiert und welche Varianten entwickelt werden. Im nächsten Schritt werden die Komponenten und Funktionen auf Grundlage der beschriebenen Anforderungen entworfen und zu einem Gesamtfahrzeug integriert. Dieser Prozess erfolgt in enger Zusammenarbeit zwischen OEM und Zulieferer. Das entwickelte Fahrzeug wird im nächsten Schritt durch Absicherungstests validiert. Das Ergebnis dieses Prozesses, der Teststatus, bildet die wesentliche Grundlage für die Entscheidung über die Freigabe des Fahrzeugs für die Produktion. Über die Freigabeentscheidung sind nur wenige Informationen öffentlich verfügbar. Im Fall der spezifikationsbasierten Absicherung (Kapitel 2.5.1) kann die Freigabe erfolgen, wenn durch die Tests nachgewiesen wird, dass alle Anforderungen erfüllt werden. Für die feldbasierte Absicherung (Kapitel 2.5.2) liegen häufig unternehmensinterne Richtlinien vor, nach denen eine Freigabe erteilt wird. Nach außen kommuniziert werden diese Vorgaben nicht. Dies ist vor dem Hintergrund möglicher rechtlicher Konsequenzen einer solchen Aussage auch nachvollziehbar. Daher lassen sich über die Abläufe dieses Prozesses nur Vermutungen anstellen. Verläuft der Produktentwicklungsprozess wie geplant, ist damit zu rechnen, dass die unternehmensinternen Vorgaben eingehalten werden und die Produktverantwortlichen die Freigabe bei Erfüllung aller Kriterien erteilen. Probleme können sich in Situationen ergeben, in denen aufgrund von kurzfristigen, ungeplanten Änderungen, die sich in der Praxis häufig nicht vermeiden lassen, bestimmte Freigabeziele nicht mehr innerhalb der Entscheidungsfrist erreicht werden können. Dann ist durch die Produktverantwortlichen eine Abwägung verschiedener Risiken durchzuführen. Auf der einen Seite steht das Risiko, dass die Änderung negative Auswirkungen auf die Produktsicherheit hat, die bisher nicht entdeckt wurden. Wird dies aufgrund der Erfahrung der Produktverantwortlichen als hoch eingestuft (z.B. anhand der Automotive Safety Integrity Level (ASIL)-Klassifizierung⁸ der geänderten Funktion), kann die Freigabe nicht erteilt werden. Auf der anderen Seite steht das finanzielle Risiko einer Start-of-Production-(SOP)-Verschiebung. Wird das Sicherheitsrisiko durch die Produktverantwortlichen als gering eingestuft, so ist davon auszugehen, dass die Freigabe erteilt wird, um eine Verschiebung des SOP zu verhindern.

⁸ Der Automotive Safety Integrity Level (ASIL) ist das Ergebnis der Gefahrenanalyse gemäß ISO 26262. Er kennzeichnet das funktionale Sicherheitsrisiko für eine zu entwickelnde Funktion (vgl. ISO 26262: Road vehicles – Functional safety, 2011).

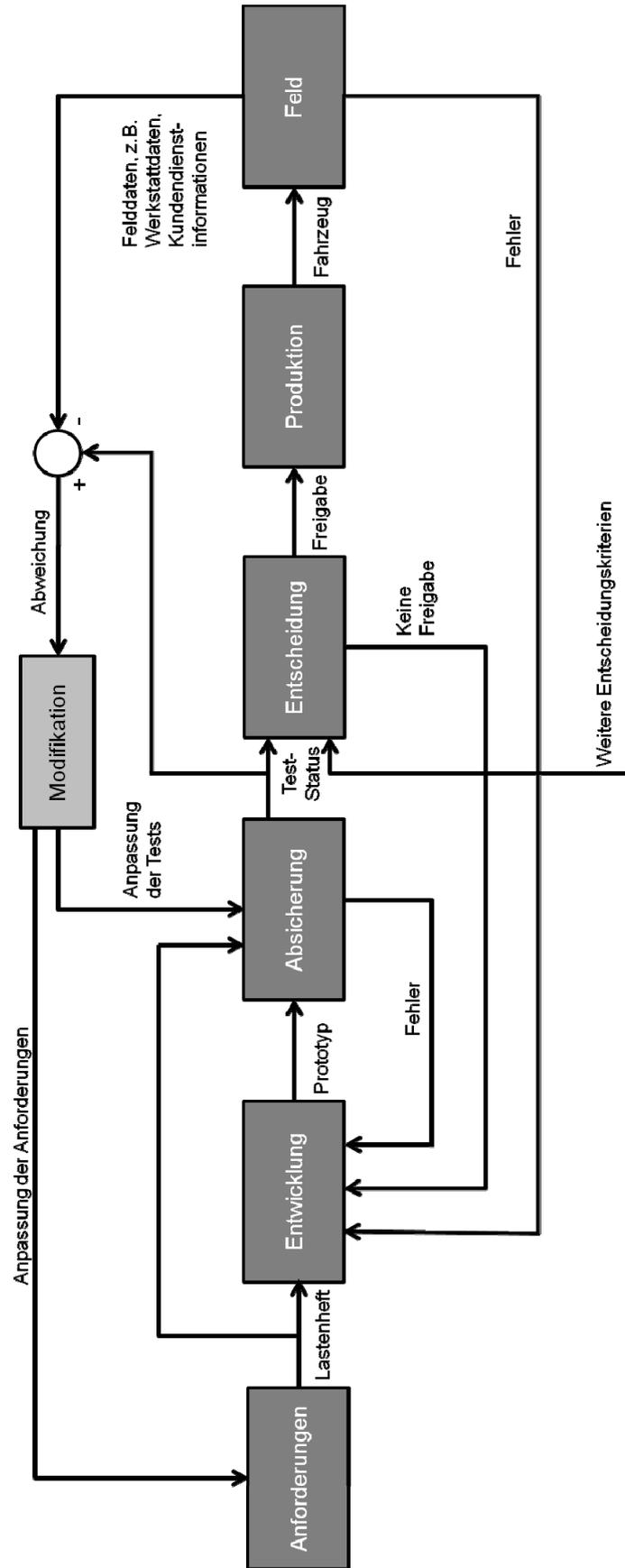


Abbildung 2-1: Rolle der Gesamtfahrzeugabsicherung im Entwicklungsprozess

In beiden Fällen ist die Entscheidung stark vom Expertenwissen und von der Erfahrung der Produktverantwortlichen abhängig, die vermutlich auch die persönlichen Konsequenzen ihrer Entscheidung (z.B. interne Karriere vs. strafrechtliche Konsequenzen) mit in den Freigabeprozess einfließen lassen. So ist anzunehmen, dass sich die subjektiven Bewertungsmetriken (z.B. bei der Klassifizierung der Änderungen hinsichtlich der Sicherheitsrelevanz) mit steigendem Projektdruck verändern. Nach der Entscheidung zur Freigabe folgt die Produktion, die anschließende Auslieferung des Fahrzeugs an den Kunden und die Nutzung im Feld. Der in Abbildung 2-1 dargestellte Entwicklungsprozess ist in der Praxis durch Rekursionsschleifen geprägt. Bauteil- oder Softwarefehler, die z.B. im Rahmen der Erprobungen gefunden werden, führen zu einer (teilweisen) Wiederholung von Entwicklungs- und Absicherungstätigkeiten. Die in der Absicherung entdeckte Fehleranzahl kann allerdings auch zur Bewertung der Absicherungsqualität herangezogen werden. Eine solche rückblickende Analyse der Absicherung bietet die Möglichkeit, die Absicherungsprozesse zu reflektieren und zu optimieren. Eine alternative Bewertung der Qualität kann z.B. anhand der Nutzung des Fahrzeugs im Feld erfolgen. Mögliche Informationen, die hierfür verwendet werden können, sind Werkstattinformationen, Kundendienstdaten, Pannenstatistiken oder Rückrufaktionen. Nachteil dieser Bewertung der Absicherung ist, dass sie erst relativ spät erfolgen kann.

2.3 Anforderungen an die Absicherung

Da die Gesamtfahrzeugabsicherung die Lücke zwischen der Produktentwicklung und dem Kundenbetrieb schließt, muss sie sicherstellen, dass das System alle Anforderungen erfüllt. Dies erklärt die hohe rechtliche Relevanz dieses Prozesses. Produkthaftungsgesetze, wie z.B. das deutsche Produkthaftungsgesetz (ProdHaftG), gewährleisten, dass sowohl Fahrzeughersteller als auch Zulieferer für ihre Produkte verantwortlich gemacht werden. Im Schadensfall ist der Nachweis zu führen, dass die Produkte nach dem Stand der Technik entwickelt wurden.⁹ Dieser repräsentiert die aktuellen Gesetze, Richtlinien und Standards sowie Patente und Veröffentlichungen.¹⁰ Daraus ergeben sich auf der einen Seite Anforderungen, die die Beschaffenheit oder Funktionalität von Produkten festlegen, z.B. die Straßenverkehrszulassungsordnung oder die EU ECE R 13 H¹¹ bzw. die FMVSS 126¹² für automobiler Bremssysteme. Auf der anderen Seite resultieren aus dem Stand der Technik aber auch Anforderungen, die die Vorgehenswei-

⁹ Reif: Automobilelektronik, 2009, S. 261

¹⁰ Reuter: Produkthaftung in Deutschland, 2011

¹¹ ECE R13 H: Uniform provisions concerning the approval of passenger cars, 2014

¹² NHTSA: FMVSS No. 126: Electronic Stability Control Systems, 2007

se bei der Entwicklung von Produkten festlegen, z.B. ISO/TS 16949¹³ und ISO/IEC 15504 (Automotive SPICE)¹⁴ oder die ISO 26262¹⁵, die die Anforderungen an die funktionale Sicherheit vorgibt. Über die durch Gesetze und Normen bedingten Anforderungen hinaus, stellen die Automobilhersteller eine große Anzahl an internen Anforderungen zur Absicherung des Gesamtfahrzeugs. Die unternehmensinternen Anforderungen werden nicht nach außen kommuniziert, wodurch lediglich Vermutungen über Umfang, Herkunft und Priorisierung angestellt werden können.

2.4 Durchführung der Absicherung

Die Gesamtfahrzeugabsicherung ist ein entscheidendes, qualitätssicherndes Element im Produktentstehungsprozess der Automobilindustrie.¹⁶ Absicherungstests sind in der Regel Black-Box-Tests, bei denen die Funktionalität eines Systems ohne Einblick in dessen innere Struktur untersucht wird.¹⁷ Sie bestehen aus einer großen Anzahl an verschiedenen Tests mit jeweils unterschiedlichen Beurteilungszielen. Typische Testkategorien für automobiler Steuergeräte nach Borgeest¹⁷ sind:

- Funktionstests
- Robustheitstests
- Benchmarks
- Recovery-Tests
- Konfigurations- und Kompatibilitätstests
- Bedienbarkeitstests
- Sicherheitstests
- Lebensdauerstests.

Gesamtfahrzeugtests setzen voraus, dass das Fahrzeug in seiner späteren Umgebung unter Betriebsbedingungen getestet wird.¹⁷ Außerdem müssen die mit dem Fahrzeug interagierenden Systeme Fahrer und Umwelt zwingend bei der Entwicklung des Test-

¹³ ISO/TS 16949: Quality management systems, 2009

¹⁴ ISO/IEC 15504: Information technology – Process assessment, 2012

¹⁵ ISO 26262: Road vehicles – Functional safety, 2011

¹⁶ Sundmark et al.: An Exploratory Case Study of Testing, 2011

¹⁷ Borgeest: Elektronik in der Fahrzeugtechnik, 2010, S. 227 f.

designs mit berücksichtigt werden.¹⁸ Daher erfolgt ein großer Teil der Erprobungen im realen Fahrversuch. Dies ist mit erheblichem Personalaufwand sowie hohen Kosten für die Ausstattung der Fahrzeuge und die Benutzung der Teststrecken etc. verbunden woraus sich insgesamt ein enormer Absicherungsaufwand ergibt. Die Testschwerpunkte, die dabei überprüft werden, sind u.a. Fahrdynamik, Akustik, Referenz und Benchmark, Testen von Assistenzsystemen und Fahrbarkeitsanalysen.¹⁹ Eine große Herausforderung bei der Absicherung im realen Fahrversuch ist die Objektivierung der Bewertung sowie die Herstellung reproduzierbarer Bedingungen.¹⁹ Dies ist die Voraussetzung für die Sicherstellung einer gleichbleibenden Qualität des Endproduktes.

Ein weiterer Testbereich sind Hardware-in-the-Loop (HiL)-Tests. Dort wird das entwickelte System in einem simulierten Fahrzeugumfeld erprobt.²⁰ Die hohe Reproduzierbarkeit von Fahrmanövern sowie die Möglichkeit der Testautomatisierung sind Vorteile dieses Absicherungsansatzes gegenüber den Tests im Realfahrzeug.²⁰ Darüber hinaus ist die einfache Verfügbarkeit von verschiedenen Umgebungsbedingungen sowie Fahrzeugparametern vorteilhaft.²⁰ Kvasnicka et al.²¹ und Mao et al.²² zeigen Beispiele für die Anwendung von HiL-Simulationen in der Absicherung von Bremssystemen. Darüber hinaus beschreiben Holzmann et al.²³ einen Computer Aided Engineering (CAE)-basierten Homologationsansatz für Electronic Stability Control (ESC)-Systeme. Die Qualität der HiL-Tests hängt maßgeblich von den zugrundeliegenden Datensätzen ab (z.B. mathematisches Fahrzeugmodell, Fahrzeugparameter). Eine hohe Detailtiefe erfordert in der Regel einen großen Modellierungsaufwand.²⁴ Außerdem können bisher nicht alle Bewertungskriterien (z.B. Fahrkomfort, Fahrbarkeit, Bedienbarkeit) objektiv gemessen werden. Aus diesem Grund ist der Nutzen von HiL-Tests für die Gesamtfahrzeugabsicherung heute noch begrenzt.

In der Automobilindustrie existieren für diese rechtlich relevante, letzte Absicherung weder standardisierte Vorgehensweisen noch klare Bewertungsmaßstäbe.²⁵ Dies macht einen Vergleich unterschiedlicher Absicherungsstrategien schwierig. Ein Ansatz zur

¹⁸ Düser: X-in-the-Loop, 2010, S. 31

¹⁹ Düser: X-in-the-Loop, 2010, S. 45 f.

²⁰ Borgeest: Elektronik in der Fahrzeugtechnik, 2010, S. 228 f.

²¹ Kvasnicka et al.: Durchgängige Simulationsumgebung, 2006

²² Mao et al.: CAE supported ESC development/release process, 2012

²³ Holzmann et al.: Simulationsbasierte ESP-Homologation für Pkw, 2012

²⁴ Langermann: Beitrag zur durchgängigen Simulationsunterstützung, 2008

²⁵ Sundmark et al.: An Exploratory Case Study of Testing, 2011

Bewertung von Absicherungskonzepten, der diesem Problem begegnet, wird von Singer et al.²⁶ vorgestellt.

2.5 Paradigmen der Gesamtfahrzeugabsicherung

In der Gesamtfahrzeugabsicherung finden sich zwei grundsätzliche Denkweisen, die nachfolgend dargestellt werden.

2.5.1 Paradigma 1: Spezifikationsbasierte Absicherung

Das erste Paradigma stellt die spezifikationsbasierte Absicherung dar. Diese beruht auf einer Testfallgenerierung, die auf einer vollständigen und korrekten Spezifikation aufbaut. Ein „idealer“ V-Modell-Prozess²⁷ dient als hinreichende Grundlage des Paradigmas. Hierbei werden zunächst die vollständigen Produkteigenschaften spezifiziert, z.B. in einem Anforderungsbaum. Ein Anforderungsbaum stellt eine hierarchische Sortierung der Spezifikation auf den einzelnen Abstraktionsebenen (z.B. Produkthanforderungen, funktionale Anforderungen, technische Anforderungen, Implementierungsanforderungen) dar.²⁸ Eine wichtige Voraussetzung bei der Erstellung der Anforderungsdefinition ist die lückenlose und eindeutige Rückverfolgbarkeit der Spezifikation. Diese stellt sicher, dass jede Produkthanforderung auf Fahrzeugebene mit Anforderungen auf unteren Systemebenen, die diese konkretisieren, verknüpft wird. Eine solche Dokumentation dient als Basis für den rechten Ast des V-Modells, d.h. die Verifikation und Validierung der Produkteigenschaften.²⁸ Die Grundlage für die Absicherung bildet die Ableitung von Testfällen aus den Anforderungen. Um eine belastbare Verifikation durchführen zu können, wird jeder Anforderung mindestens ein Testfall zugeordnet. Durch die geforderte Vollständigkeit der Spezifikation wird dabei sichergestellt, dass das Fahrzeug umfassend geprüft wird. Kann durch die Absicherung nachgewiesen werden, dass alle Anforderungen durch die Implementierung korrekt umgesetzt werden, ist die Entwicklung beendet. Die abschließende Freigabeentscheidung resultiert aus den zugrundeliegenden Testergebnissen.

²⁶ Singer et al.: Gesamtfahrzeugabsicherung, 2015

²⁷ Für Details zum V-Modell wird auf Schäuffele, Zurawka: Automotive Software Engineering, 2003, S. 22 ff. verwiesen.

²⁸ Wilhelm et al.: Funktionale Sicherheit und ISO 26262, 2015, S. 96

2.5.2 Paradigma 2: Feldbasierte Absicherung

Das zweite Paradigma ist die feldbasierte Absicherung. Sie stützt sich auf eine Abbildung der Nutzung des Fahrzeugs im Betrieb. Dazu wird die Absicherung derart durchgeführt, dass ein dem Nutzungsprofil des Fahrzeugs repräsentierendes Kollektiv von Fahrsituationen überprüft wird. Die Absicherung ist damit unabhängig vom vorausgegangenen Entwicklungsprozess. Ein entscheidender Faktor für die Qualität der Absicherung ist die Auswahl der Tests. Das Situationskollektiv kann systematisch aus Feldstudien (z.B. die „Naturalistic Driving Study“, siehe Neale et al.²⁹) oder statistischen Daten (siehe z.B. Weitzel³⁰) ermittelt werden. Alternativ ist auch eine empirische Testfallgenerierung möglich, die auf dem Expertenwissen der Tester und der Erfahrung vorangegangener Absicherungen aufbaut. Die feldbasierte Absicherung kann als Prozess der vorgezogenen Feldbeobachtung aufgefasst werden. Dieser Auslegung des Begriffs folgend, ist die Absicherung ein Prognosewerkzeug, das die Eigenschaften des Fahrzeugs in der Zukunft, d.h. im Kundenbetrieb, prädiziert.

2.5.3 Vergleich der Paradigmen

Das erste Paradigma der Gesamtfahrzeugabsicherung setzt eine vollständige und eindeutig nachvollziehbare Anforderungsdefinition als Basis für die Testfallgenerierung voraus. Dieses Absicherungskonzept wird z.B. in der ISO 26262 für Sicherheitsanforderungen explizit gefordert. Erste Umsetzungen sind daher in der Automobilindustrie schon zu finden. Da dieser Ansatz jedoch nicht nur zur Sicherstellung der Funktionalen Sicherheit erforderlich ist, sondern auch allgemein der Gewährleistung einer hohen Produktqualität dient, kann die spezifikationsbasierte Absicherung grundsätzlich auf alle Teile des Gesamtsystems Fahrzeug übertragen werden. Es gibt daher auch die Bestrebung, dieses Konzept für alle Entwicklungsbereiche zu übernehmen.³¹ Der Ansatz bringt einige Vorteile mit sich. So ist der Aufwand für die Testfallgenerierung selbst niedrig. Die Tests können direkt aus den Anforderungen abgeleitet werden. Die benötigte vollständige Anforderungsdefinition kann darüber hinaus auch in anderen Situationen hilfreich sein. So erleichtert eine derartige Spezifikation die Kommunikation zwischen den Entwicklungspartnern (OEM ↔ Systemzulieferer ↔ Komponentenzulieferer). Eindeutig beschriebene Anforderungen liefern klare Entwicklungsziele und definierte Testkriterien. Es sind demnach weniger Abstimmungen und weniger Rekursionsschleifen in der Entwicklung zu erwarten. Außerdem kann eine derartige Spezifikation auch

²⁹ Neale et al.: The 100 Car Naturalistic Driving Study, 2002

³⁰ Weitzel: Objektive Bewertung der Kontrollierbarkeit, 2013, S. 37 ff.

³¹ Wilhelm et al.: Funktionale Sicherheit und ISO 26262, 2015, S. 99

den Umgang mit Änderungen im Entwicklungsprozess verbessern. So können mit Hilfe der Anforderungsdefinition die Auswirkungen der Änderung bestimmt (siehe z.B. Mokammel³²) und der benötigte Wiederholungstestbedarf ermittelt werden. Ein weiterer Vorteil ergibt sich in der an die Absicherung anschließende Freigabe. Die Anforderungen liefern klare Kriterien für die Freigabe. Wenn alle Tests bestanden sind und damit alle Anforderungen in zuvor definiertem Maß erfüllt sind, kann freigegeben werden. Die Entscheidung ist damit objektivierbar, was insbesondere aufgrund der rechtlichen Relevanz ein nicht zu vernachlässigender Vorteil ist. Das Paradigma besitzt jedoch auch Nachteile. Die Qualität der Absicherung ist bei diesem Konzept stark von der Qualität der Spezifikation abhängig. Anforderungen werden heute aber häufig formlos und in natürlicher Sprache dokumentiert, was eine automatische Testfallermittlung erschwert.³³ Darüber hinaus sind sie oft unvollständig oder mehrdeutig formuliert.³³ Die Ermittlung einer vollständigen Anforderungsdefinition ist zudem sehr aufwändig und stellt für komplexe Systeme wie ein Gesamtfahrzeug eine große Herausforderung dar.³⁴ So müssen alle geforderten Fahrzeugeigenschaften objektiv beschreibbar sein. Im Bereich der Fahrdynamik ist dies z.B. nicht immer der Fall. Hier wird das „Fahrgefühl“ häufig subjektiv durch Experten bewertet. Auch die Evaluierung der Kontrollierbarkeit beruht zumeist auf Expertenwissen.³⁵ Selbst bei einer vollständigen Spezifikation ist allerdings nicht automatisch sichergestellt, dass die daraus resultierenden Absicherungstests ausreichend sind. Es fehlen derzeit noch Methoden, die die Feldabdeckung bei der Ermittlung der Anforderungen mit berücksichtigen. Forschungsbedarf besteht außerdem noch darin, wie sichergestellt werden kann, dass die Testfälle aus den Anforderungen richtig abgeleitet werden. Ebenso ist offen, wie aus einer Menge an Testfällen zur Überprüfung einer Anforderung der richtige Testfall ausgewählt wird. Die Umsetzung dieses Paradigmas wird damit heute zwar angestrebt, ist aber aufgrund der o.g. Hindernisse noch nicht Stand der Technik in der Gesamtfahrzeugabsicherung.

Folglich wird bei heutigen Absicherungsprozessen auf das zweite Paradigma zurückgegriffen. Im Allgemeinen wird dabei die empirische Testfallgenerierung angewandt, bei der die einzelnen Testfälle durch einen Ingenieur³⁶ entwickelt und spezifiziert werden. Diese manuelle Testfallgenerierung hat nach Siegl et al.³⁷ viele Nachteile. So erweist

³² Mokammel: Impact Analysis of Graph-based Requirements Models, 2013

³³ Siegl et al.: Formal Specification and Systematic Model-Driven Testing, 2011

³⁴ Wilhelm et al.: Funktionale Sicherheit und ISO 26262, 2015, S. 99

³⁵ Siehe z.B. Weitzel: Objektive Bewertung der Kontrollierbarkeit, 2013, S. 17 ff.; Neukum, Krüger: Fahrerreaktionen bei Lenksystemstörungen, 2003

³⁶ Der Begriff Ingenieur wird in dieser Arbeit geschlechtsneutral verwendet und umfasst sowohl Ingenieurinnen als auch Ingenieure.

³⁷ Siegl et al.: Formal Specification and Systematic Model-Driven Testing, 2011

sich z.B. die Bestimmung der Testabdeckung als schwierig. Überdies lässt sich nicht vermeiden, dass ein oder mehrere wichtige Testszenarien unentdeckt bleiben. Nur in kleinen Absicherungsteilbereichen wird eine systematische Ableitung von Testsituationen durchgeführt (siehe z.B. Weitzel³⁸). Die repräsentative Feldabdeckung wird heute nicht formal nachgewiesen, es liegt vielmehr ein „gewachsenes“ Testkollektiv vor, das sich durch seine Anwendung als Stand der Technik bestätigt. Vorteil dieses Vorgehens ist, dass auf bestehende Testfälle zurückgegriffen werden kann und lediglich bei neuen Funktionen bzw. Komponenten Veränderungen der Testkataloge erforderlich werden. Außerdem decken die Testfälle eine Vielzahl an früher aufgetretenen Problemen ab, an die die Ingenieure vermutlich erst nicht denken. Der Aufwand für die Testfallgenerierung ist dementsprechend gering. Die Prozesse der Entwicklung und der Absicherung sind relativ losgelöst voneinander. Dies hat den Vorteil geringer Abhängigkeiten, birgt allerdings den Nachteil, dass Synergien nicht genutzt werden können. So wird die Absicherung durch die Abwesenheit von eindeutigen Bestehenskriterien für Testfälle erschwert. Die Qualität der Absicherung ist bei diesem Paradigma stark von der Qualität des empirisch ermittelten Testkollektivs abhängig. Es gibt derzeit keine Bewertungsmaßstäbe, mit denen diese Qualität bestimmt werden kann. Dies führt insbesondere bei der an die Gesamtfahrzeugabsicherung anschließenden Freigabe zu Schwierigkeiten, da keine eindeutigen Kriterien ermittelt werden können, nach denen beurteilt wird. Die Freigabeentscheidung orientiert sich am Stand der Technik. So werden die Fahrzeuge z.B. einer Dauerläuferprobung unterzogen, die der Sicherstellung eines stimmigen und in der Summe funktionierenden Gesamtkonzeptes für alle Märkte unter Berücksichtigung verschiedener Fahrzeugkonfigurationen dient.³⁹ Dazu wird das Fahrzeug unter extremen dynamischen, klimatischen und topographischen Bedingungen auf unterschiedlichen Streckenprofilen und mit verschiedenen Belastungszuständen getestet. Die Detektionsrate für Fehler ist dabei abhängig von der Länge der Dauerläuferprobung. Für funktionale Fehler ergibt sich zu Beginn eine hohe Detektionsrate, die mit der Dauer der Erprobung stark abfällt. Aus dem Erwartungswert der Verteilung der entdeckten Fehler über die Testdauer kann dann ein Kriterium für die Freigabe abgeleitet werden (z.B. Streckenlänge in km). Die Festlegung des konkreten Grenzwertes (z.B. 100.000 km oder 120.000 km) unterliegt dabei jedoch trotz Expertenwissen einer subjektiven Willkür. Dieses Beispiel zeigt, dass die Freigabeentscheidung im Falle der feldbasierten Absicherung am Stand der Technik orientiert ist, die Interpretation des Standes der Technik sowie die Festlegung der Anforderungen an die Freigabe (d.h. Grenzwerte) jedoch subjektiven Einflüssen unterliegen. Die Freigabeentscheidung ist demnach stark vom Expertenwissen und der Erfahrung der Produktverantwortlichen abhängig. Bei

³⁸ Weitzel: Objektive Bewertung der Kontrollierbarkeit, 2013, S. 33

³⁹ Klem et al.: Erprobung der Langzeitqualität des neuen Audi A4, 2007, S. 64

inkrementell erfolgenden Entwicklungsprozessen mag dies noch akzeptabel sein. Fraglich ist jedoch die Eignung dieses Vorgehens bei ganz neuen Herausforderungen wie veränderten Antriebskonzepten (z.B. Brennstoffzelle), innovativen Fahrwerken (z.B. breiten- und längenverstellbare Achsen) oder dem automatisierten Fahren.

Die oben diskutierten Paradigmen bilden die Grundlage für die Entwicklung einer Testauswahlstrategie zur Absicherung von Änderungen auf Gesamtfahrzeugebene. Eine Testauswahlmethode, die allein auf dem Konzept der spezifikationsbasierten Absicherung aufbaut, ist aufgrund der zuvor beschriebenen, fehlenden Voraussetzungen derzeit nicht ausreichend für die Absicherung von Änderungen. Sofern diese Prämissen nicht gegeben sind, ist daher die feldbasierte Absicherung unverzichtbar. Die Testauswahl hat daher zwingend feldbasierte Absicherungsaspekte zu berücksichtigen.

2.6 Herausforderungen der Gesamtfahrzeugabsicherung

Die Gesamtfahrzeugabsicherung unterliegt diversen Herausforderungen, die in den vorangegangenen Abschnitten identifiziert und in diesem Abschnitt zusammengefasst werden.

Die hohe rechtliche Relevanz kennzeichnet die erste Herausforderung. Durch die Gesamtfahrzeugabsicherung ist sicherzustellen, dass alle Anforderungen an das Fahrzeug, die aus Gesetzen und Normen resultieren, erfüllt werden. Die Qualität der Absicherung ist daher ein entscheidendes Bewertungskriterien für neue Konzepte in der Gesamtfahrzeugabsicherung.

Die zweite Herausforderung stellt die hohe Systemkomplexität des Fahrzeuggesamtsystems dar, die aus der großen Anzahl an individuellen Systemen (Hardware und Software) sowie aus der intensiven Konnektivität zwischen diesen Systemen resultiert. Sie führt dazu, dass ein vollständiges Testen aller Kombinationen der Eingangsparameter einer Funktion unter ökonomisch vertretbarem Aufwand nicht mehr realisierbar ist.⁴⁰ Einer Testauswahlstrategie, die diese Komplexität berücksichtigt, kommt demnach eine zentrale Bedeutung für die Absicherungsqualität zu.

Der hohe Aufwand für die Durchführung der Absicherungsprozesse ist die dritte Herausforderung. Er ist zum einen Folge der großen Varianten- und Ausstattungsvielfalt heutiger Fahrzeuge. Zum anderen ist er darin begründet, dass der Hauptanteil der Absicherungstests im realen Fahrversuch erfolgt. Die Ausgaben für Verifikations- und Vali-

⁴⁰ Nörenberg: Effizienter Regressionstest von E/E-Systemen nach ISO 26262, 2012, S. 26

dierungsaktivitäten bilden einen großen Anteil an den Gesamtentwicklungskosten.⁴¹ Diese beinhalten die Aufwände für die Fahrzeuge und Komponenten, Teststrecken, Simulationsumgebungen, Personalkosten usw. Es ist daher anzustreben, diesen Aufwand mit neuen Konzepten zu reduzieren oder zumindest nicht zu vergrößern.

Die vierte Herausforderung stellt der späte Zeitpunkt der Gesamtfahrzeugabsicherung im Entwicklungsprozess dar. Die Absicherung kann erst durchgeführt werden, wenn alle Komponenten und Systeme verfügbar sind.⁴² Dies stellt hohe Anforderungen an die Effizienz der Prozesse.

Änderungen betreffen die fünfte Herausforderung der Gesamtfahrzeugabsicherung. Sie sind besonders kritisch, wenn sie erst spät in den Prozess einfließen (siehe Sundmark et al.⁴²), weil sie zum Teil durch frühere, nicht bestandene Tests notwendig wurden. Außerdem erfordern sie eine (teilweise) Wiederholung der Gesamtfahrzeugabsicherung, so dass die begrenzt zur Verfügung stehenden Ressourcen stärker belastet werden. Der Umgang mit Änderungen ist somit ein kritischer Erfolgsfaktor. Die Identifikation der Effekte der Änderungen sowie die Bestimmung des resultierenden Wiederholungstestbedarfs sind daher ein wichtiges Forschungsfeld.

Der beschriebenen letzten Herausforderung widmen sich die nachfolgenden Kapitel dieser Arbeit, in denen Testauswahlmethoden zur Absicherung von Änderungen entwickelt werden. Die dargestellten anderen vier Herausforderungen stellen Randbedingungen dar, die beim Entwurf und der Evaluierung der Testauswahlmethoden zu berücksichtigen sind.

⁴¹ Albers: Vorwort zu Band 47 der Forschungsberichte des Instituts für Produktentwicklung, 2010

⁴² Sundmark et al.: An Exploratory Case Study on Testing, 2011

3 Absicherung von Änderungen in der Automobilindustrie

Der Umgang mit Änderungen ist ein kritischer Erfolgsfaktor in der Gesamtfahrzeugabsicherung. In diesem Kapitel wird beschrieben, wie das Änderungsmanagement in heutigen Entwicklungsprozessen gestaltet wird. Darüber hinaus werden unterschiedliche Konzepte zur Auswahl von Testfällen bei Änderungen vorgestellt und analysiert. Am Ende des Kapitels folgt eine Diskussion der Herausforderungen heutiger Änderungsprozesse.

3.1 Änderungen

Änderungen sind nach Zanner et al.⁴³ alle nachträglichen Anpassungen von freigegebenen, d.h. verbindlich festgelegten Arbeitsergebnissen. Im Rahmen dieser Arbeit werden ausschließlich technische Änderungen analysiert und die Auswirkungen auf das Produkt ermittelt. Für diese Betrachtungsweise werden Änderungen daher als alle nachträglichen Anpassungen eines verbindlich festgelegten Produktstandes definiert.

Änderungen bilden einen festen Bestandteil aller Entwicklungsprozesse, da die Produktentwicklung immer in iterativen Schritten erfolgt.⁴⁴ Sie gehen nach Gustavsson⁴⁵ aus funktionalen Erweiterungen, funktionalen Modifikationen, Kostenreduktionsmaßnahmen oder Adaptionen durch neue Hardware hervor. Eckert et al.⁴⁶ unterscheiden zwischen auslösenden („initiated“) und auftauchenden („emergent“) Änderungen. Auslösende Änderungen entspringen einer äußeren Ursache. Viele dieser Änderungen sind zu Beginn eines Entwicklungsprozesses bekannt, z.B. Kunden- oder Marktanforderungen, neue oder modifizierte gesetzliche Anforderungen für Produkte oder Prozesse, Innovationen (neue Materialien, Komponenten oder Software) oder Probleme mit vergangenen Produkten. Manche Änderungen dieser Kategorie werden aber auch erst im Laufe des Entwicklungsprozesses durch neue Kundenanforderungen oder gegenwärtige technische Fortschritte bekannt. Auftauchende Änderungen sind dem aktuellen Stand

⁴³ Zanner et al.: Änderungsmanagement bei verteilten Standorten, 2002

⁴⁴ Milling, Jürging: Der Serienanlauf in der Automobilindustrie, 2008, S. 69

⁴⁵ Gustavsson: Architecting Automotive Product Lines, 2010

⁴⁶ Eckert et al.: Change and customization in complex engineering domains, 2004

des Entwurfs geschuldet. Sie resultieren aus Problemen mit dem konkreten Produkt. Beispiele sind Abweichungen von Kundenanforderungen, mangelnde Funktionsfähigkeit des Produkts oder Sicherheitsrisiken. Auftauchende Änderungen treten zu allen Zeitpunkten im Entwicklungsprozess und auf allen Integrationsstufen auf (z.B. im Entwurf, Test, Prototypenstadium, Produktion oder Betrieb).

Änderungen verursachen in der Entwicklungs- und Konstruktionsphase bis zu 40 % der Gesamtaufwände.⁴⁷ Zu den unmittelbaren Kosten zählen nach Lindemann und Reichwald⁴⁸ der Arbeitsaufwand für die Änderungsplanung und –verwaltung, die Nacharbeit, die Verschrottungskosten sowie Werkzeugänderungskosten und die Entwertung von vorhandenen Beständen. Je später eine Änderung im Entwicklungsprozess eintritt, desto aufwändiger ist ihre Behebung. Dafür gibt es zwei Gründe. Auf der einen Seite wird der Prozess immer kritischer, je weniger Zeit für die Integration der Änderung zur Verfügung steht. Daher werden mehr Ressourcen benötigt, um mit der Änderung in der verbleibenden Zeit umzugehen. Auf der anderen Seite ist das Produkt zu einem fortgeschrittenen Entwicklungszeitpunkt schon stark integriert, so dass der Einfluss der Änderung größer ist und mehr Nacharbeit und –prüfung notwendig werden.⁴⁹

Aufgrund dieser Zusammenhänge existieren in der Automobilindustrie viele Strategien, die das Ziel verfolgen, nachträgliche Änderungen zu vermeiden. Beispiele dafür finden sich vor allem im Qualitätsmanagement. Eine häufig verwendete Methode ist die Failure Modes and Effects Analysis (FMEA). Sie identifiziert zunächst potenzielle Fehler und bestimmt die Auswirkungen auf das Gesamtsystem oder die zugrundeliegenden Prozesse. Anschließend wird der Fehler anhand einer Risikoprioritätszahl bewertet und ggf. Maßnahmen definiert.⁵⁰ Dadurch können potenzielle Fehler frühzeitig entdeckt werden, so dass mögliche nachträgliche Änderungen vermieden und somit Entwicklungszeit und Entwicklungskosten reduziert werden.^{51 52} Eine weitere Qualitätstechnik ist das Quality Function Deployment (QFD). Die QFD-Methode beschreibt einen systematischen Ansatz zur schrittweisen und strukturierten Umsetzung von Kundenanforderungen in messbare Produkt- und Prozessparameter.⁵³ Mithilfe des sog. House of Quality, das eine Korrelationsmatrix darstellt, die Kunden- und Designanforderungen gegenüberstellt, werden Produktspezifikationen vermieden, die nicht den

⁴⁷ Lindemann, Reichwald: Integriertes Änderungsmanagement, 1998

⁴⁸ Lindemann, Reichwald: Integriertes Änderungsmanagement, 1998, S. 35

⁴⁹ Eckert et al.: Change and customization in complex engineering domains, 2004

⁵⁰ Struss, Fraracci: Automated model-based FMEA of a braking system, 2012

⁵¹ Eversheim et al.: Simultaneous Engineering, 1995, S. 76 ff.

⁵² Dickmann: Schlanker Materialfluss, 2009, S. 51

⁵³ Westkämper: Null-Fehler-Produktion in Prozessketten, 1997, S. 31 f.

Kundenerwartungen entsprechen, so dass die gesamte Produktentstehungsphase auf eine maximale Kundenorientierung ausgerichtet werden kann.⁵⁴

Trotz des Einsatzes dieser Fehlerpräventionsmethoden treten späte Änderungen in der Praxis häufig auf.⁵⁵ Studien zeigen, dass etwa 40 % der Änderungen erst ab der Nullserienphase durchgeführt werden.⁵⁵ Da späte, ungeplante Änderungen besonders kritisch sind, befasst sich diese Arbeit mit Änderungen in der Gesamtfahrzeugabsicherung, die den letzten Schritt im Entwicklungsprozess darstellt, bevor das Produkt gefertigt und an den Kunden ausgeliefert wird.

3.2 Änderungsmanagement

Das Änderungsmanagement definiert die Tätigkeiten und Prozesse, mit denen Änderungen strukturiert und nachvollziehbar in den Produktentstehungsprozess integriert werden.⁵⁶ Zu den Aufgaben des Änderungsmanagements zählen laut DIN 69901-5⁵⁷ die Erfassung, Bewertung, Entscheidung, Dokumentation und Steuerung der Umsetzung von Änderungen. Eine Änderung wird in der Automobilindustrie als kleines Projekt innerhalb des Gesamtentwicklungsprojektes betrachtet.⁵⁸

Gustavsson⁵⁹ beschreibt den Änderungsmanagementprozess in fünf Schritten (siehe Abbildung 3-1). Zunächst wird die Änderung durch eine Bedarfsanalyse identifiziert. Darauf folgend werden die Auswirkungen der Änderung ermittelt. Unterschiedliche Lösungsmöglichkeiten werden im nächsten Schritt aufgestellt. Auf deren Grundlage wird daraufhin eine Entscheidung über die Änderung getroffen. Zuletzt wird die Änderung dann implementiert und validiert. Dazu muss der Absicherungsprozess erneut durchlaufen werden. Eine zentrale Frage dabei ist, welche Testfälle erneut durchgeführt werden müssen. Diesem Aspekt widmet sich diese Arbeit.

Aufgrund der Häufigkeit von Änderungen ist die effiziente Gestaltung des Änderungsprozesses von hoher Bedeutung für den Erfolg von Innovationsprojekten.⁶⁰ So kann das Änderungsmanagement eine schnelle und flexible Reaktion auf neue Kundenwünsche

⁵⁴ Wannenwetsch: Integrierte Materialwirtschaft und Logistik, 2004, S. 164

⁵⁵ Milling, Jürging: Der Serienanlauf in der Automobilindustrie, 2008, S. 71

⁵⁶ Grande: 100 Minuten Konfigurationsmanagement, 2013, S. 79

⁵⁷ DIN 69901: Projektmanagement, 2009

⁵⁸ Borgeest: Elektronik in der Fahrzeugtechnik, 2010, S. 263

⁵⁹ Gustavsson: Architecting Automotive Product Lines, 2010

⁶⁰ Schuh: Innovationsmanagement, 2012, S. 222

und geänderte Marktanforderungen ermöglichen, die Belastung eines Unternehmens aufgrund geringerer Kosten pro Änderung verringern, eine höhere Innovationsrate erzielen und sich positiv auf den Aufbau von Wissen und das Sammeln von Erfahrung auswirken. Angesichts der Herausforderung, immer komplexere Produkte in immer kürzeren Zeitabständen und unter wachsendem Kostendruck zu entwickeln, kommt der Weiterentwicklung der Methoden des Änderungsmanagements eine zentrale Rolle zu.

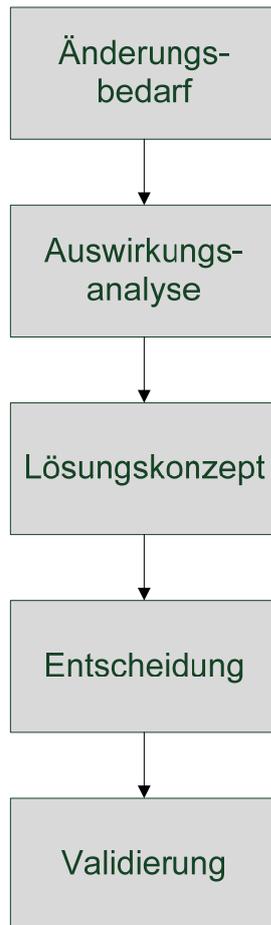


Abbildung 3-1: Änderungsmanagement nach Gustavsson

3.3 Anforderungen an das Änderungsmanagement

Bei der Entwicklung neuer Konzepte für den Umgang mit Änderungen sind die Anforderungen an das Änderungsmanagement in der Automobilindustrie, die sich aus Gesetzen und Normen ergeben, zu berücksichtigen.

Das Softwareänderungsmanagement wird in Automotive SPICE (Software Process Improvement and Capability Determination)⁶¹ geregelt. Automotive SPICE ist ein Standard zur Durchführung von Bewertungen von Unternehmensprozessen mit dem Schwerpunkt auf der Softwareentwicklung. Automotive SPICE bestimmt für jeden einzelnen Prozess im Entwicklungsablauf einen Reifegrad. Die Teilreifegrade gehen dann in den Gesamtreifegrad ein, so dass ein detaillierter Überblick über die Stärken und Verbesserungspotenziale des untersuchten Projektes entsteht. Der in Automotive SPICE spezifizierte Änderungsmanagementprozess (SUP.10)⁶² ist darauf ausgerichtet sicherzustellen, dass Änderungen kontrolliert, verfolgt und beobachtet werden. Dazu wird ein Änderungsmanagementplan etabliert, der die Änderungsmanagementstrategie des Unternehmens enthält. Dort werden die Änderungsmanagementaktivitäten, wie z.B. die Identifikation, Dokumentation, Analyse und Implementierung der Änderungen, definiert. Das zentrale Element des Änderungsmanagementprozesses ist der sog. Change Request (CR), der verwendet wird, um eine einzelne Änderung nachzuverfolgen. Er spezifiziert das Ziel der Änderung, identifiziert die Effekte für die existierenden Systeme und enthält die Verantwortlichkeiten sowie den Status und die Kritikalität der Änderung. Automotive SPICE fordert eine Überprüfung der Änderung, um sicherzustellen, dass sie den gewünschten Effekt hat und die jeweiligen Ziele und Verifikationskriterien erfüllt werden. Eine erneute Durchführung von Absicherungstests wird bei Automotive SPICE darüber hinaus explizit bei Systemintegrationstests gefordert (ENG.7)⁶³. Hier wird bei Änderungen von Softwarebausteinen (einschließlich der damit verbundenen Anforderungen sowie des damit verbundenen Designs und Codes) eine Regressionsstrategie zur erneuten Integration und Verifikation verlangt. Unter einem Regressionstest wird dabei ein „erneuter Test eines bereits getesteten Objektes nach dessen Modifikation mit dem Ziel nachzuweisen, dass durch die vorgenommenen Änderungen keine Defekte eingebaut oder bisher maskierte Defekte freigelegt wurden“ verstanden.⁶⁴

⁶¹ Bei Automotive SPICE handelt es sich um eine Anpassung des internationalen Standards ISO/IEC 15504 (Information technology – Process assessment) an die Belange der Automobilindustrie.

⁶² VDA: Automotive SPICE Prozessassessment, 2007, S. 94 ff.

⁶³ VDA: Automotive SPICE Prozessassessment, 2007, S. 70 ff.

⁶⁴ VDA: Automotive SPICE Prozessassessment, 2007, S. 182

Anforderungen an das Änderungsmanagement in der Automobilindustrie resultieren auch aus der ISO 26262⁶⁵, die sich mit der funktionalen Sicherheit von Kraftfahrzeugen auseinandersetzt. ISO 26262 (Part 8, Clause^o7) beschreibt den erforderlichen Änderungsmanagementprozess. Er enthält folgende Schritte: Änderungsanfrage, Änderungsanalyse, Entscheidung über die Änderung und Änderungsimplementierung und -dokumentation. Das zentrale Element ist die Analyse der Auswirkungen der Änderungen auf die funktionale Sicherheit des Systems („impact analysis“). In Abhängigkeit von den Ergebnissen dieser Analyse wird der Sicherheitslebenszyklus angepasst und die betroffenen Arbeitsprodukte überarbeitet. Darüber hinaus ist eine erneute Verifikation und Validierung der betroffenen Systembestandteile erforderlich. Dabei ist zu berücksichtigen, dass im Falle von verschiedenen Konfigurationen bzw. Kalibrierungsdatensätzen die notwendigen Verifizierungsmaßnahmen für die funktionale Sicherheit für alle Varianten auf System- und Fahrzeugebene nachgewiesen werden müssen (ISO 26262-4, 6.4.8).

Sowohl Automotive SPICE als auch die ISO 26262 fordern demnach eine erneute Absicherung des Gesamtsystems nach der Integration von Änderungen. Wie genau diese durchzuführen ist, bleibt aber offen. Die Standards lassen damit Freiraum in der Auswahl der Technik und Methoden. Die Herausforderung bei der Entwicklung einer Testauswahlmethode für die Absicherung von Änderungen besteht also darin, innerhalb des von der Norm gegebenen Spielraums eine zugeschnittene Lösung zu erarbeiten.

3.4 Testauswahlmethoden zur Absicherung von Änderungen

Änderungen in Komponenten oder Softwarefunktionen können eine direkte oder indirekte Auswirkung auf andere Komponenten oder Funktionen haben. Es existieren zahlreiche Ansätze, die sich mit der Identifikation dieser Effekte und der Bestimmung des benötigten Wiederholungstestbedarfs auseinandersetzen. Eine Auswahl wird im folgenden Abschnitt beschrieben.

3.4.1 Manuelle Testauswahl

In heutigen Testprozessen werden die einzelnen Testfälle überwiegend durch einen Ingenieur aus einer Menge von bereits existierenden Tests ausgewählt.⁶⁶ Die Auswahl

⁶⁵ ISO 26262: Road vehicles – Functional safety, 2011

⁶⁶ Siegl et al.: Formal Specification and Systematic Model-Driven Testing, 2011

kann statisch, z.B. auf Grundlage einer Bewertung von Risiko und Wichtigkeit einer Funktion, oder dynamisch, basierend auf den Auswirkungen der Änderung, erfolgen.⁶⁷ In vielen Fällen werden auch statische und dynamische Ansätze kombiniert, indem Testfälle priorisiert werden, von denen dann innerhalb des gegebenen Zeit- und Ressourcenrahmens möglichst viele durchgeführt werden.⁶⁷ Bei den dynamischen Konzepten werden idealerweise die Tests ausgewählt, welche die Teile des Systems prüfen, die modifiziert wurden oder die von der Änderung außerdem betroffen sind.⁶⁸ Die größte Herausforderung besteht dabei im Management des Wissens über die Abhängigkeiten des Systems.⁶⁸ Aufgrund der zunehmenden Vernetzung heutiger Fahrzeugsysteme sind auch die Effekte von Änderungen immer größer. Änderungen können sich daher potenziell auf eine große Anzahl von anderen Komponenten auswirken. Außerdem existieren viele indirekte Abhängigkeiten zwischen Modulen, die nur durch ein detailliertes Fachwissen über die Modulfunktionalitäten und ihre Implementierungen aufgedeckt werden können.⁶⁸

Die Qualität der manuellen Testauswahl ist daher stark von der individuellen Erfahrung und Systemkenntnis des Entscheiders abhängig.⁶⁹ Häufig führt dies zu Problemen, wenn die Entwicklung eines Systems und der Test desselben nicht durch dasselbe Team durchgeführt werden. In diesem Fall sind die Tester oft nicht ausreichend vertraut mit der Systemimplementierung, so dass sie nicht alle Wechselwirkungen und indirekten Abhängigkeiten der Änderungen überblicken. Dementsprechend ist die Auswahl von relevanten Testfällen sehr schwierig.⁷⁰ Probleme können auch entstehen, wenn erfahrene Entwickler das Unternehmen verlassen und damit die Begründung für bestimmte Design- und Implementierungsentscheidungen verloren geht. Ohne dieses Wissen können Änderungen zu unerwünschten Nebeneffekten führen.⁷⁰

Die manuelle Testauswahl hat mehrere Vorteile. Zunächst kann das Expertenwissen hier gut eingebracht werden.⁷¹ Darüber hinaus sind keine aufwändigen Systeme oder Verfahren notwendig, die erst durchlaufen werden müssen. So kann der Test schnell begonnen werden.⁷² Diesen Vorteilen stehen aber auch erhebliche Nachteile gegenüber. Die Testauswahl im Fall von Änderungen ist eine schwierige Entscheidung. Systemtests sind häufig auf der Ebene von Fahrzeugeigenschaften definiert, während die Änderungen Komponenten oder Software betreffen. Das detaillierte Zusammenspiel zwischen Tests

⁶⁷ Engström et al.: Improving Regression Testing Transparency and Efficiency, 2011

⁶⁸ Buchgeher: Towards Tool-Support for Test Case Selection, 2013

⁶⁹ Engström et al.: Improving Regression Testing Transparency and Efficiency, 2011

⁷⁰ Buchgeher: Towards Tool-Support for Test Case Selection, 2013

⁷¹ Paulic et al.: Verbesserter Softwaretest für Steuergeräte, 2012, S. 151

⁷² Alt: Car-Multimedia Systeme Modell-basiert testen mit SysML, 2009, S. 29

und Änderung ist von Hand fast unmöglich zu ermitteln. Die manuelle Testauswahl beinhaltet daher immer einen bestimmten Anteil von Unsicherheit.⁷³ Damit besteht die Gefahr, dass notwendige Tests unentdeckt bleiben und so unvollständige Testsuiten ausgewählt werden.⁷⁴ Darüber hinaus ist die Reproduzierbarkeit der Entscheidung nicht gewährleistet.⁷⁵ Bei der manuellen Testauswahl besteht zudem die Schwierigkeit, einen Überblick über die Testabdeckung zu erhalten und damit die Qualität der ausgewählten Testsuite zu beurteilen.⁷⁶ Außerdem kann der personelle und zeitliche Aufwand für die Testauswahl bei großen Systemen mit einer hohen Gesamtanzahl an Testfällen schnell sehr hoch werden.⁷⁷

3.4.2 Automatisierte Testauswahl

Neben der manuellen Testauswahl, die im Wesentlichen auf dem Wissen und der Erfahrung des Testers basiert, sind viele Konzepte beschrieben, die eine automatisierte Testauswahl durchführen. Die Auswahl der Testfälle erfolgt dabei auf Grundlage einer Analyse von unterschiedlichen Informationen (z.B. Softwarecode, Systemarchitekturen, Spezifikationen, statistische Daten).⁷⁸ Es kann ferner zwischen voll- und teilautomatisierten Verfahren unterschieden werden, wobei vollautomatisierte Konzepte ohne den Eingriff einer Testperson auskommen und teilautomatisierte Ansätze partiell auf Expertenwissen zurückgreifen.⁷⁹

Änderungsnachverfolgung auf Funktions- und Komponentenebene

In der Literatur sind zahlreiche Konzepte beschrieben, die die Auswirkungen von Änderungen auf Funktions- und Komponentenebene untersuchen.

Einige Ansätze verwenden die Design Structure Matrix (DSM). Die Idee geht auf Steward⁸⁰ und Eppinger et al.⁸¹ zurück. Die DSM ist eine Methodik, die die Erfassung, Modellierung, Analyse und Synthese (in Grenzen) von Verbindungen von Elementen in Systemnetzwerken ermöglicht. Das zugrundeliegende Modell bildet eine quadratische

⁷³ Juergens et al.: Regression Test Selection of Manual System Tests in Practice, 2011

⁷⁴ Paulic et al.: Verbesserter Softwaretest für Steuergeräte, 2012, S. 151

⁷⁵ Alt: Car-Multimedia Systeme Modell-basiert testen mit SysML, 2009, S. 29

⁷⁶ Paulic et al.: Verbesserter Softwaretest für Steuergeräte, 2012, S. 151

⁷⁷ Alt: Car-Multimedia Systeme Modell-basiert testen mit SysML, 2009, S. 29

⁷⁸ Buchgeher: Towards Tool-Support for Test Case Selection, 2013

⁷⁹ Alt: Car-Multimedia Systeme Modell-basiert testen mit SysML, 2009, S. 29

⁸⁰ Steward: The design structure system, 1981

⁸¹ Eppinger et al.: A model-based method for organizing task in product development, 1994

Matrix, bei der die Systemelemente in den Zeilen und Spalten der Matrix eingetragen werden. Die Abhängigkeiten zwischen den Elementen werden durch die Zellen der Matrix repräsentiert. Jede Zelle zeigt dann eine numerische oder binäre Repräsentation der Verbindung zwischen dem Element der Spalte und dem Element der Zeile. Die Effekte von Änderungen können dann durch Restrukturierung der Abhängigkeitsmatrix bestimmt werden.^{82 83}

Beispiele für Änderungsnachverfolgungsmethoden, die auf dieser Technik aufbauen, sind die Component–Funktion Propagation Method⁸⁴ und die Change Propagation Method (CPM)⁸⁵. Die Component–Function Propagation Method analysiert die Abhängigkeiten zwischen Systemelementen auf einer binären Basis (0 – keine Verbindung, 1 – physikalische oder funktionale Verbindung), wohingegen die CPM für jede Abhängigkeit zwischen zwei Elementen das Risiko bestimmt, inwieweit das eine Element von einer Änderung des anderen Elements betroffen ist. Das Risiko wird als Kombination der Wahrscheinlichkeit der Abhängigkeit sowie der Auswirkungsstärke der Änderung bewertet. Die Zuordnung konkreter Zahlenwerte für die Risiken erfolgt dabei durch eine Expertenbewertung.

Cohen et al.⁸⁶ zeigen einen ähnlichen Ansatz, die sog. Change Favorable Representation (C-FAR), die eine Änderungsnachverfolgung auf Attributebene darstellt. Das C-FAR Produktmodell besteht aus Elementen, Attributen, die diese Elemente beschreiben, und Beziehungen, die die Verbindungen zwischen Elementen repräsentieren. Elemente werden bei C-FAR als Vektoren dargestellt, wobei die Dimension jedes Vektors mit der Anzahl seiner Attribute korrespondiert. Die verschiedenen Elemente werden durch Matrizen miteinander verbunden, wobei jedes Matrixelement als Verbindungswert bezeichnet wird. Sie zeigen quantitativ, wie eine Änderung eines Attributs eines Elements das entsprechende Attribut des anderen Elements beeinflusst. Die Verbindungswerte werden dabei in drei Kategorien eingeteilt: H („high linkage“), M („medium linkage“) und L („low linkage“). Um die Konsequenzen einer Änderung von einem Ausgangselement zu einem Zielelement zu bestimmen, wird ein Einflusspfad berechnet, der aus einer Reihe von Vektor- und Matrixmultiplikationen besteht.

⁸² Clarkson et al.: Prediction change propagation in complex design, 2001

⁸³ DSM Web: Design Structure Matrix, 2014

⁸⁴ Flanagan et al.: A functional analysis of change propagation, 2003

⁸⁵ Clarkson et al.: Predicting change propagation in complex design, 2001; Keller et al.: Visualizing change propagation, 2005; Jarratt et al.: Development of a product model to support engineering change management, 2004

⁸⁶ Cohen et al.: C-FAR, change favorable representation, 2000

Raffaelli et al.⁸⁷ führen eine Änderungsanalyse auf Grundlage eines Modells der Produktarchitektur durch. Die Hauptfunktionen eines Systems werden dabei durch Blöcke („black-boxes“) repräsentiert, die über Input-Output-Verbindungen miteinander verknüpft werden. Die Verbindungen können in drei Kategorien eingeteilt werden: Material-, Energie- und Signalströme. Darüber hinaus findet eine Abbildung von Funktionen und physikalischen Komponenten statt, bei der die Komponenten detailliert modelliert werden (z.B. geometrische Abmessungen, Materialien, Farbe) und über physikalische Verbindungen (z.B. Schweißnaht) oder konzeptionelle Abhängigkeiten (z.B. Position) miteinander verkettet werden.

Ein anderer Ansatz, der auf einem Produktmodell aufbaut, wird von Eckert et al.⁸⁸ beschrieben. Ein Produkt wird dabei durch drei Typen von Elementen dargestellt: direkte Parameter, funktionale Parameter und Verhaltensparameter. Physikalische Komponenten werden durch direkte Parameter repräsentiert (z.B. Geometrie, Material, Gewicht, Leistung). Funktionen resultieren aus der Interaktion direkter Parameter. Sie werden in gewünschte Parameter und Nebeneffektparameter (z.B. Geräusch, Vibration, elektromagnetische Verträglichkeit (EMV)) unterteilt. Das Verhalten eines Produktes kann dann aus den Interaktionen der Funktionen bestimmt werden. Mit Hilfe dieses Modells können so die Effekte von Änderungen in direkten Parametern auf das Verhalten des Systems ermittelt werden.

Die vorgestellten Konzepte für die Nachverfolgung von Änderungen auf Funktions- und Komponentenebene haben gemeinsam, dass sie die Effekte von Änderungen auf Basis eines vereinfachten Systemmodells bestimmen. Einige verwenden dazu mathematische Modelle, während andere Ansätze auf komplexere Produktmodelle mit detailliertem Einblick auf Eigenschaften, Funktionen und Verhalten zurückgreifen. Ein großer Vorteil dieser Konzepte ist, dass sie so allgemein sind, dass sie grundsätzlich auf automobiler Systeme auf Fahrzeuggesamtebene anwendbar sind. Die Ergebnisse der Änderungsanalyse sind allerdings stark von der Qualität des zugrundeliegenden Systemmodells abhängig. Darüber hinaus kann der Aufwand für die Aufstellung eines Modells für große Systeme mit vielen Verbindungen sehr hoch werden, wenn dieses Modell nicht automatisch im Entwicklungsprozess erzeugt wird, sondern zusätzlich zu gängigen Produktmodellen (z.B. Computer Aided Design (CAD)-Modell oder Fahrdynamiksimulationsmodell) generiert werden muss.

Änderungseinflussanalyse auf Softwareebene

In der Softwaretechnik existieren zahlreiche Konzepte zur Bestimmung der Auswirkungen von Änderungen. Die Ansätze werden unter dem Begriff Change Impact Analysen

⁸⁷ Raffaelli et al.: Development of a multilayer change propagation tool for modular products, 2007

⁸⁸ Eckert et al.: Change customization in complex engineering domains, 2004

(CIA) zusammengefasst. CIAs zielen darauf ab, die Teile eines Softwaresystems zu identifizieren, die von der Änderung betroffen sind. Auf Grundlage einer Anzahl von definierten Änderungen („change set“), werden die Elemente bestimmt, die durch diese Änderungen beeinflusst werden und zusätzliche Modifikation benötigen („impact set“).⁸⁹

Lehnert⁹⁰ präsentiert einen Überblick über CIA-Methoden. Dazu analysiert er 150 verschiedene Literaturquellen. Er unterscheidet drei Bereiche von Änderungsanalysen.

Ansätze des ersten Bereichs untersuchen die Auswirkungen von Änderungen durch den Rückschluss aus dem Quelltext. Sie analysieren Vererbungsbeziehungen, Methodenauf-rufverhalten oder andere Abhängigkeiten zwischen Programmelementen. Statische Ansätze evaluieren dabei Aufrufgraphen, Slices⁹¹ und andere Repräsentationen des Quelltextes während dynamische Konzepte oder Online-Methoden Ausführungspfade untersuchen.

Der zweite Anwendungsbereich von CIAs sind formale Modelle. Sie werden weiter unterteilt in Architektur- und Anforderungsmodelle. Architekturmodelle sind z.B. UML (Unified Modeling Language)-Komponentendiagramme⁹², die Systeme, Subsysteme, Komponenten und Klassen von Softwaresystemen illustrieren. UML-Modelle erlauben die Bestimmung der Auswirkungen von Änderungen auf einer abstrakteren Ebene als Quelltext. Auch wenn Anforderungen in einer formalen Modellierungssprache umgesetzt werden, können sie für die Abschätzung der Auswirkungen von Änderungen herangezogen werden.

Die dritte Gruppe von CIAs basiert auf sonstigen Entwicklungsartefakten, z.B. Dokumentationen, Fehlertrackern oder Konfigurationsdateien. Kombinationen der verschiedenen Typen von CIAs sind ebenso möglich.

In der Softwareentwicklung existiert eine Vielzahl von Konzepten, die sich mit der Identifikation der Auswirkungen von Änderungen befassen. Einige Methoden benötigen spezifische Informationen, wie z.B. Quelltext. Diese sind in automobilen Entwicklungsprojekten nicht immer verfügbar. So hat ein Fahrzeughersteller i.d.R. keinen Zugriff auf den Quelltext eines Bremsregelsteuergeräts, das beim Zulieferer entwickelt wurde. Andere Ansätze nutzen weniger spezifische Informationsgrundlagen, wie z.B. UML-Modelle. Diese Konzepte sind grundsätzlich auf Systeme auf Fahrzeugebene

⁸⁹ Yaszdanshenas, Moonen: Fine-grained change impact analysis, 2012

⁹⁰ Lehnert: A review of change software change impact analysis, 2011

⁹¹ Ein Slice ist ein Teil eines Programmes, der nur die Anweisungen enthält, die das Verhalten eines bestimmten Programmpunktes (sog. Slicing-Kriterium) beeinflussen (Backward Slice) oder von ihm beeinflusst werden (Forward Slice).

⁹² OMG: Unified Modeling Language, 2014

anwendbar. Allerdings ist auch hier der Aufwand für die Bereitstellung und Wartung dieser Modelle hoch.

Testauswahlmethoden für den Regressionstest

Ein weiteres Konzept für die Reduktion des Absicherungsaufwandes im Fall von Änderungen ist der Regressionstest. Regressionstests zielen darauf ab, ein bereits überprüftes Testobjekt nach seiner Modifikation erneut zu evaluieren. Sie beabsichtigen, zu bestätigen, dass die Implementierung der Änderung zu keinen weiteren Fehlern geführt hat. Studien schätzen, dass Regressionstests ca. 80 % des Gesamttestaufwandes von Softwareprojekten ausmachen.⁹³ Der Testaufwand kann reduziert werden, wenn nur eine Auswahl von allen Testfällen erneut durchgeführt wird.⁹⁴ Dazu werden Testauswahlmethoden (sog. regression test selection techniques (RTS)) angewandt. Studien zeigen die Effizienz dieses Konzeptes.⁹⁵

RTS-Methoden nutzen unterschiedliche Informationen als Grundlagen für die Ermittlung des Wiederholungstestbedarfs. Viele Konzepte verwenden Repräsentationen des Quelltextes, um die Teile der Software zu identifizieren, die von den Änderungen betroffen sind und ordnen diesen Bereichen Testfälle zu. Beispiele für diesen Ansatz beschreiben Vokolos und Frankl⁹⁶, Rothermel und Harrold⁹⁷ oder Gallagher et al.⁹⁸. Komponenten-basierte RTS-Techniken basieren auf Softwareelementen, für die kein Quelltext verfügbar ist. Sie analysieren die Schnittstellen der Komponenten oder nutzen Aufrufgraphen als Grundlagen für die Auswirkungsanalyse. Zheng et al.⁹⁹ zeigen einen beispielhaften Ansatz. Andere Konzepte^{100, 101, 102} verwenden Architekturmodelle (z.B. UML-Diagramme), die mit Testfällen verbunden sind, um die Effekte von Änderungen zu bestimmen und den benötigten Wiederholungstestbedarf zu berechnen. Anforderungen sind eine andere Quelle für RTS-Techniken. Gorthi et al.¹⁰³ beschreiben bspw. einen

⁹³ Kaner: Improving the maintainability of automated test suites, 1997

⁹⁴ Kim et al.: An empirical study of regression test application frequency, 2005

⁹⁵ Leung, White: A study of integration testing and software regression, 2005; Rothermel et al.: Empirical studies of test-suite reduction, 2002; Khan et al.: The impact of test case reduction, 2009

⁹⁶ Vokolos, Frankl: Pythia: A regression test selection tool, 1997

⁹⁷ Rothermel, Harrold: Empirical studies of safe regression test selection techniques, 1998

⁹⁸ Gallagher et al.: Reducing regression test size by exclusion, 2007

⁹⁹ Zheng et al.: An initial study of a lightweight process, 2005

¹⁰⁰ Zhao et al.: Change impact analysis to support architectural evolution, 2002

¹⁰¹ Muccini et al.: Software architecture-based regression testing, 2006

¹⁰² Briand et al.: Automating regression test selection, 2009

¹⁰³ Gorthi et al.: Specification-based approach to select regression test suite, 2008

spezifikationsbasierten Ansatz, der auf UML-Aktivitätsdiagrammen aufbaut. Ein anderes Beispiel für diesen Ansatz ist Chittimalli und Harrold's Requirement RTS¹⁰⁴, die Anforderungen mit Quelltext verkettet und Testfälle durch die Analysen von Testpfaden auswählt.

Caliebe et al.¹⁰⁵ und Nörenberg¹⁰⁶ haben das Konzept des Regressionstests von Softwareanwendungen auf eingebettete Systeme übertragen. Caliebe et al.'s Regressionstestmethodik nutzt ein Black-Box-Systemmodell, das sog. Component-Dependency-Model (CDM). Es wird aus der Systemarchitektur (z.B. AUTOSAR-Architektur¹⁰⁷) sowie den zugehörigen Anforderungen abgeleitet und in eine Graphstruktur überführt, um damit Pfadanalysen durchführen zu können. Mit Graphalgorithmen aus der Softwaretechnik können die Effekte von Änderungen berechnet und der benötigte Wiederholungstestbedarf ermittelt werden. Nörenberg beschreibt einen spezifikationsbasierten Ansatz für Regressionstests. Er verwendet das Konzept der „Funktionsorientierung“ (FO) von Daimler als Basis für die Analyse der Änderungsauswirkung.

Alle beschriebenen Regressionstestmethodiken benötigen eine detaillierte Repräsentation des zu analysierenden Systems für die Bestimmung der Effekte der Änderung. Einige Regressionstesttechniken sind sehr spezialisiert, weil sie z.B. einen bestimmten Typ von Quelltext in einer bestimmten Programmiersprache benötigen. Andere Methoden nutzen eine weniger spezifische Informationsbasis, wie z.B. UML-Modelle. Diese Konzepte sind grundsätzlich auch auf Fahrzeugebene anwendbar.

Statistische Versuchsplanung

Die statistische Versuchsplanung stellt eine weitere Alternative für die Testauswahl in Absicherungsprozessen dar. Das Ziel der statistischen Versuchsplanung ist die Realisierung eines gewünschten experimentellen Ergebnisses mit minimalem Aufwand. Durch die gleichzeitige Veränderung mehrerer Faktoren, können die Einflüsse verschiedener Faktoren auf einen bestimmten Parameter oder eine Parametergruppe bestimmt werden.¹⁰⁸

Ungermann¹⁰⁹ zeigt einen Ansatz, der die statistische Versuchsplanung nutzt, um den Stichprobenumfang von Zuverlässigkeitstests in automobilen Absicherungsprozessen zu reduzieren. Darüber hinaus erlaubt das Konzept die systematische Bestimmung des

¹⁰⁴ Chittimalli, Harrold: Regression test selection on system requirements, 2008

¹⁰⁵ Caliebe et al.: Dependency-based test case selection and prioritization, 2012

¹⁰⁶ Nörenberg: Effizienter Regressionstest von E/E-Systemen nach ISO 26262, 2012

¹⁰⁷ Siehe z.B. AUTOSAR: <http://autosar.org>, 2014

¹⁰⁸ Borgeest: Elektronik in der Fahrzeugtechnik, 2008

¹⁰⁹ Ungermann: Zuverlässigkeitsnachweis und Zuverlässigkeitsentwicklung, 2009

benötigten Testaufwandes im Fall von späten Änderungen. Die Grundlage für die Versuchsplanung bildet eine Analyse des komponentenspezifischen Ausfallverhaltens auf verschiedenen Versuchsstrecken. Außerdem werden Komplexitätsklassen der Komponenten und der Reifegrad des Projektes mit berücksichtigt. Durch die Integration von Informationen über die Nutzung des Fahrzeugs beim Kunden kann daraus ein modellspezifischer Testplanungsstandard entwickelt werden.

Burgdorf¹¹⁰ nutzt ebenfalls die statistische Versuchsplanung für die Bestimmung von Testaufwand in automobilen Absicherungsprozessen von E/E-Systemen. Auf der Basis eines Vorhersagemodells für zukünftige Bordnetzkonfigurationen, werden repräsentative Fahrzeugkonfigurationen für Absicherungstests identifiziert. Das Konzept wird erweitert durch eine kundenrelevante Risikodefinition, die verwendet wird, um die statistische Risikoreduktion zu berechnen, die mit den verschiedenen Fahrzeugtestkonfigurationen erreicht werden kann. Durch eine iterative Optimierung wird dann ein Testplan ermittelt.

Die statistische Versuchsplanung erlaubt eine zielgerichtete Auswahl von Testfällen. Vorteil dieses Ansatzes ist, dass er keine umfangreiche Systemkenntnis voraussetzt, da er auf statistischen Daten aufbaut. Dies ermöglicht der Testperson Schwerpunkte auf fehleranfällige oder risikobehaftete Tests zu legen. Andererseits benötigen statistische Ansätze eine große Datenbasis, um verlässliche Aussagen treffen zu können. Des Weiteren führt die statistische Natur der Verfahren zu Testdefiziten in Bereichen, die statistisch unterrepräsentiert sind.

Zusammenfassung

Die beschriebenen Methoden der Änderungsauswirkungsanalyse und der Testauswahl lassen sich hinsichtlich ihrer zugrundeliegenden Informationsbasis unterscheiden (siehe Abbildung 3-2).

Die erste Gruppe nutzt den Softwarequelltext als Informationsbasis. Beispiele für diese Kategorie sind CIAs. Sie erlauben eine detaillierte Analyse der Auswirkungen einer Änderung kombiniert mit einer zuverlässigen Testauswahl. Im Fall von großen, stark vernetzten Systemen ist dieser Ansatz jedoch aufgrund des hohen Berechnungsaufwandes zeitlich nicht effizient einsetzbar, so dass diese Methoden lediglich auf kleine, abgegrenzte Systeme anwendbar sind. Darüber hinaus ist das Konzept auf spezifische Testobjekte beschränkt (z.B. Quelltext in C++) und daher nicht allgemein auf Fahrzeugebene anwendbar.

¹¹⁰ Burgdorf: Eine kunden- und lebenszyklusorientierte Produktfamilienabsicherung, 2010

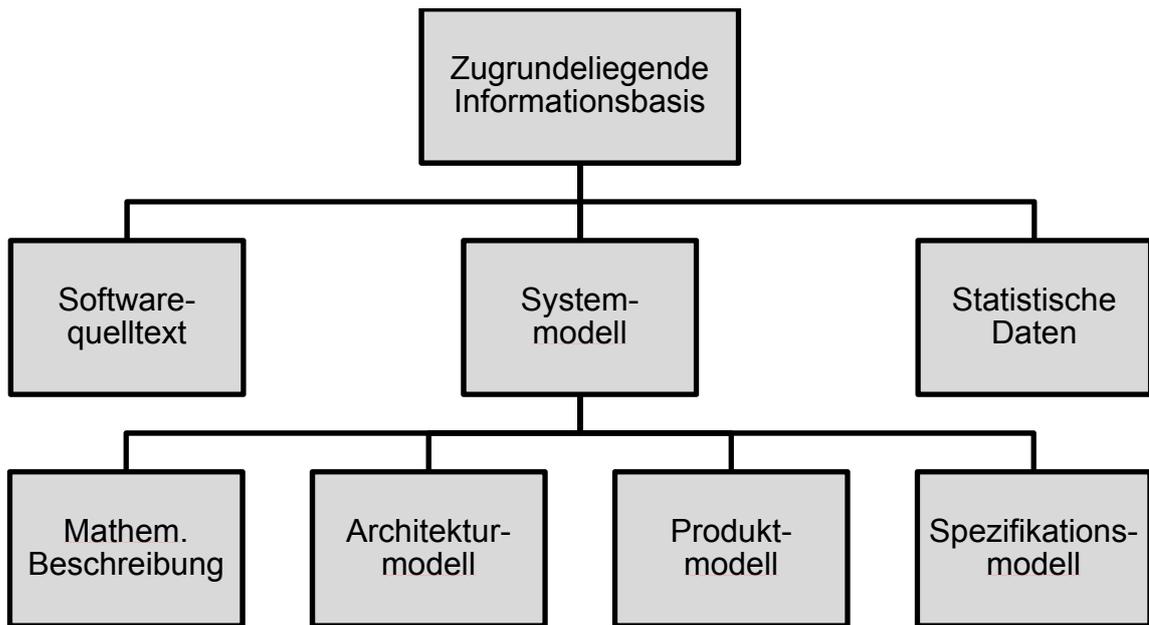


Abbildung 3-2: Klassifikation der Informationsgrundlage für Testauswahlmethoden

Die zweite Gruppe bilden Systemmodelle. Diese Klasse kann weiter unterschieden werden in mathematische Systembeschreibungen (z.B. Matrizen oder Vektordarstellungen wie bspw. CPM), Architekturmodelle (z.B. UML-Modelle), Produktmodelle (z.B. CDM) und Spezifikationsmodelle (z.B. FO). Alle Ansätze dieser Kategorie sind grundsätzlich auf Fahrzeugebene einsetzbar. Sie erlauben die Bestimmung der Auswirkungen von Änderungen in der Software genauso wie auf physikalischer Ebene. Nachteilig wirkt sich jedoch der hohe Aufwand für die Erstellung, Wartung und Verwaltung der Modelle aus, insbesondere bei großen Systemen mit vielen Wechselwirkungen. Teilweise kann dieser Aufwand durch automatische Modellgeneratoren reduziert werden. Das Ergebnis der Einflussanalyse ist dann aber immer noch stark von der Qualität der zugrundeliegenden Modellierung abhängig. Jede Interaktion, die nicht modelliert ist, führt dabei zu Defiziten in der Testabdeckung.

Die letzte Gruppe von Ansätzen nutzt statistische Daten (z.B. statistische Versuchsplanung). Diese Ansätze ermöglichen eine zielgerichtete Auswahl von Testfällen, ohne detailliertes Systemwissen zu erfordern. Nachteilig ist jedoch, dass Wechselwirkungen zwischen Komponenten oder Funktionen, die statistisch niedrig sind, vernachlässigt werden, so dass diese Testauswahl nicht sicher i.S.v. Vollständigkeit ist.

3.5 Herausforderungen bei der Absicherung von Änderungen

Nachfolgend werden die Herausforderungen des automobilen Änderungsmanagements zusammenfassend beschrieben.

Die erste Herausforderung ist die Änderungshäufigkeit. Änderungen treten zu allen Zeitpunkten im Entwicklungsprozess auf und können alle Systembestandteile (Hardware- und Software) betreffen.

Der Aufwand für die Umsetzung von Änderungen stellt die zweite Herausforderung dar. Dabei ist zu berücksichtigen, dass die Behebung einer Änderung in der Regel umso aufwändiger ist, je später sie im Entwicklungsprozess eintritt. Kurzfristige Änderungen in der Gesamtfahrzeugabsicherung, die den letzten Entwicklungsschritt vor der Produktion und anschließenden Auslieferung an den Kunden darstellt, sind damit besonders kritisch. Eine hohe Effizienz des Änderungsmanagements ist daher anzustreben.

Die dritte Herausforderung betrifft die Abwesenheit von Vorschriften oder Normen zu konkreten Methoden des Änderungsmanagements. Der Umgang mit Änderungen in der Automobilindustrie erfolgt zwar strukturiert auf Grundlage eines definierten Änderungsmanagementprozesses, der aus diversen Normen (z.B. Automotive SPICE und ISO 26262) abgeleitet ist. Konkrete Vorschläge, wie einzelne Methoden umzusetzen sind (z.B. Regressionsteststrategie), werden in den Vorschriften allerdings nicht genannt. Dies führt dazu, dass in der Praxis unternehmensindividuelle Lösungen vorherrschen. Da diese i.d.R. nicht nach außen kommuniziert werden, kann die Ergebnisqualität nur schwer verglichen werden.

Eine weitere Herausforderung ist die Absicherung von Änderungen. Sie ist erforderlich, um nachzuweisen, dass die Sicherheit und Qualität des Fahrzeugs nach Implementierung der Änderung erhalten bleibt. Um nicht die gesamte Testsuite erneut prüfen zu müssen, ist eine Testauswahl notwendig, d.h. die Identifizierung der Tests aus der Gesamtmenge aller Absicherungstests, die für die Überprüfung der Änderung notwendig sind. Die Testauswahl wird insbesondere durch die hohe Systemkomplexität beeinträchtigt, die die vollständige Ermittlung der Auswirkungen einer Änderung erschwert. In der Praxis herrscht derzeit die manuelle Testauswahl vor, deren Qualität stark von dem Expertenwissen und der Erfahrung der Testperson abhängig ist. In Anbetracht der Wichtigkeit der Testauswahl ist es daher notwendig, zu untersuchen, wie Ingenieure dabei besser unterstützt werden können.

3.6 Präzisierung der Ausgangslage, der Aufgabenstellung und des Vorgehens

Die Gesamtfahrzeugabsicherung ist ein wichtiger, qualitätssichernder Schritt im Produktentstehungsprozess in der Automobilindustrie, da sie die Lücke zwischen der Entwicklung und der Produktion bzw. dem Nutzen des Fahrzeugs im Kundenbetrieb schließt. Herausforderungen für die Gesamtfahrzeugabsicherung ergeben sich aus der hohen rechtlichen Relevanz, der großen Systemkomplexität des Fahrzeugesamtsystems, dem hohen Aufwand für die Durchführung der Absicherungsprozesse und der kurzen zur Verfügung stehenden Zeit. Ein kritischer Erfolgsfaktor in der Gesamtfahrzeugabsicherung ist der schnelle und effiziente Umgang mit Änderungen, die regelmäßig auftreten. Änderungen erfordern eine (teilweise) Wiederholung der Absicherungstests. Eine zentrale Herausforderung dabei ist die Testauswahl, d.h. die Identifizierung der Tests aus der Gesamtmenge aller Absicherungstests auf Gesamtfahrzeugebene, die für die Überprüfung der Änderung notwendig sind. Damit wird sichergestellt, dass die Änderung korrekt implementiert und keine weiteren Fehler durch sie integriert wurden.

Die Analyse des Standes der Technik zeigt, dass im industriellen Umfeld die manuelle Testauswahl vorherrscht. Ansätze für eine automatisierte, systematische Testauswahl sind in der Literatur beschrieben, allerdings in der Praxis eher wenig verbreitet. Dies führt dazu, dass die Qualität der Testauswahl derzeit stark von dem Expertenwissen und der Erfahrung der Testperson abhängig ist. Es ist darüber hinaus sehr wahrscheinlich, dass die Testperson überflüssige Testfälle in die Testsuite integrieren, um Sicherheitsrisiken zu minimieren, so dass das Testen von Änderungen unnötigerweise kostspielig ist. In Anbetracht der Wichtigkeit der Gesamtfahrzeugabsicherung ist es daher notwendig zu untersuchen, wie Ingenieure bei der Testauswahl besser unterstützt werden können.

Forschungsbedarf besteht daher in der Entwicklung von systematischen und transparenten Strategien für die Auswahl von Testfällen bei Änderungen in der Gesamtfahrzeugabsicherung. Die Forschungsfrage, die sich daraus ergibt, lautet:

Existiert eine Testauswahlmethode, die eine Identifikation der erforderlichen Tests im Fall von Änderungen in der Gesamtfahrzeugabsicherung ermöglicht?

Eine Testauswahlmethode, die auf die Bestimmung des benötigten Wiederholungstestbedarfs im Fall von Änderungen in der Gesamtfahrzeugabsicherung anwendbar ist, muss diverse Anforderungen erfüllen. Aufgrund der Tatsache, dass es sich bei der Gesamtfahrzeugabsicherung um die letzte Stufe im Testprozess handelt, ist dieser Prozessschritt von hoher rechtlicher Relevanz (vgl. Abschnitt 2.3). Die Ergebnisse der Testaus-

wahl müssen daher sicher sein, d.h. es dürfen keine notwendigen Tests vernachlässigt werden. Da der Absicherungsaufwand allgemein sehr hoch ist (vgl. Abschnitt 2.6), sollte die Testauswahlmethode zudem so wenige Tests wie möglich selektieren. Um praktisch anwendbar zu sein, ist es außerdem notwendig, dass die Testauswahlmethode effizient durchführbar ist (vgl. Abschnitt 3.5). Sie sollte daher möglichst auf bereits existierende Informationen zurückgreifen und in vorhandene Prozesse integriert werden können. Da die Gesamtfahrzeugabsicherung das Automobil auf höchster Systemebene betrachtet (vgl. Abschnitt 2.4), ist es außerdem obligatorisch, dass die Auswahlmethode eine hohe Generalität in Form von Adaption auf verschiedene Systeme und unterschiedliche Arten von Änderungen besitzt. Daneben ist zu beachten, dass die Methode innerhalb der derzeit vorherrschenden, feldbasierte Absicherung anwendbar ist (vgl. Abschnitt 2.5.3).

Ziel der folgenden Ausführungen ist es daher, zu untersuchen, ob eine Methode existiert, die die o.g. Anforderungen erfüllt. Falls diese Frage mit „ja“ beantwortet werden kann, können im nächsten Schritt Ausgestaltungsmöglichkeiten einer solchen Testauswahl entwickelt werden. Falls diese Frage mit „nein“ beantwortet werden muss, kann eine Diskussion der Ursachen erfolgen.

Der Arbeit liegt eine deduktive Vorgehensweise zugrunde. Dabei wird das Problem der Testauswahl zur Absicherung von Änderungen auf Gesamtfahrzeugebene zunächst abstrahiert (Kapitel 4.2). Anhand des generalisierten Problems werden dann prinzipielle Lösungsmöglichkeiten aufgestellt (Kapitel 4.3) und ihre Eigenschaften verglichen (Kapitel 4.4). Im nächsten Schritt folgen eine Konkretisierung der Lösungskonzepte und eine Bewertung ihrer Eignung für die Testauswahl zur Absicherung von Änderungen auf Gesamtfahrzeugebene (Kapitel 5 und Kapitel 6). Basierend auf den Ergebnissen dieser Untersuchungen wird dann ein Gesamtkonzept für die Testauswahl aufgestellt (Kapitel 7) sowie ein Validierungskonzept entwickelt (Kapitel 8). Anschließend werden die Schlussfolgerungen für den Entwicklungsprozess diskutiert (Kapitel 9).

4 Prinzipielle Lösungsansätze für Testauswahlmethoden

In diesem Kapitel wird die Testauswahl bei Änderungen auf Gesamtfahrzeugebene zunächst abstrahiert. Darauf aufbauend werden prinzipielle Lösungsmöglichkeiten aufgestellt und verglichen.

4.1 Definition Testauswahlmethode

Unter einer Testauswahlmethode wird in Anlehnung an Franz¹¹¹ in dieser Arbeit ein regelwerkbasierendes, systematisches Vorgehen zur Auswahl von Testfällen aus einer Gesamtmenge an Absicherungstests auf Gesamtfahrzeugebene verstanden. Die Auswahl soll sicherstellen, dass eine nicht korrekte Implementierung der Änderung aufgedeckt wird und keine weiteren Fehler durch sie integriert wurden. Gleichzeitig ist der Aufwand im Vergleich zum Volltest möglichst zu reduzieren.

Ein einzelner Testfall stellt dabei eine Aktivität dar, bei der das Testobjekt, hier das Gesamtfahrzeug, mittels geeigneter Testmethoden überprüft oder ausgeführt wird.¹¹¹ Jeder Testfall verfolgt ein bestimmtes oder mehrere Testziele, wie z.B. eine gewisse Fahrsituation auszuführen oder die Übereinstimmung mit spezifischen Anforderungen zu prüfen. Er umfasst die für die Ausführung notwendigen Vorbedingungen, die Menge der Eingabewerte, die Menge der vorausgegangenen Ergebnisse und die erwarteten Nachbedingungen.¹¹¹

4.2 Abstraktion der Testauswahlsituation

Die Testsituation bei Änderungsfreigaben soll hier mathematisch mit Hilfe der Mengenlehre beschrieben werden (siehe Abbildung 4-1).

¹¹¹ Franz: Handbuch zum Testen von Web- und Mobile Apps, 2015

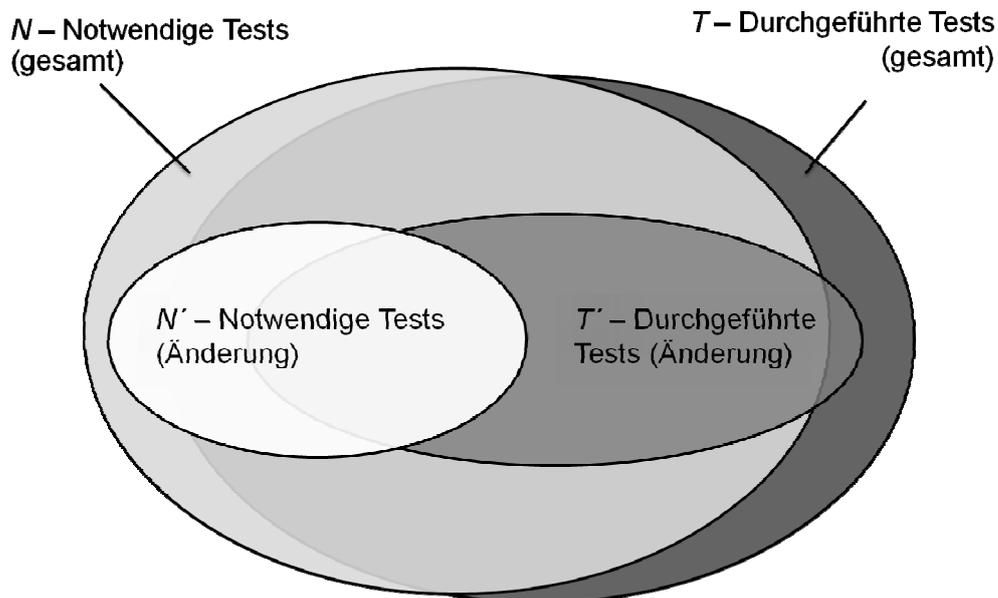


Abbildung 4-1: Testsituation bei der Absicherung von Änderungen

N repräsentiert die „wahre“ Menge der insgesamt notwendigen Tests zur Absicherung eines Systems, d.h. die Tests, die durchgeführt werden sollten, um alle internen und externen Wechselwirkungen eines Systems zu überprüfen. $|N|$, die Anzahl der Elemente aus N , ist unbekannt. Sie umfasst alle Kombinationen der erlaubten, aber auch der unerlaubten Eingangsparameter des Systems und ist daher bei komplexen Systemen wie das Gesamtfahrzeug sehr groß und daher schwer zu identifizieren.

T ist die Menge der insgesamt durchgeführten Tests zur Absicherung eines Systems. $|T|$, die Anzahl der Elemente von T , ist bekannt. Sie entspricht der Summe der Tests, die nach heutigem Stand der Technik zur Absicherung eines Systems überprüft werden. T repräsentiert z.B. die Summe aller Fahrmanöver, die von Zulieferer und Fahrzeughersteller zur Absicherung des Gesamtfahrzeugs untersucht werden.

Da die Auswirkungen von Änderungen i.d.R. nicht alle Tests betreffen, ist eine Auswahl von Testfällen für die Absicherung der Änderung durchzuführen. N' , $N' \in N$, stellt die „wahre“ Menge der notwendigen Tests zur Absicherung einer Änderung dar. T' , $T' \in T$, entspricht der Menge der durchgeführten Tests zur Absicherung einer Änderung. Dies sind z.B. die Tests, die der Testingenieur expertenbasiert aus dem Fahrmanöverkatalog auswählt oder die durch einen Auswahlalgorithmus automatisiert bestimmt werden. Dabei wird davon ausgegangen, dass zur Absicherung der Änderung keine Tests hinzugefügt werden, sondern die Absicherung auf Basis des vorhandenen

Testkatalogs erfolgt. Diese Annahme ist für alle Änderungen gültig, die Anpassungen eines verbindlich festgelegten Produktstandes umfassen.¹¹²

Die Ergebnisse der Auswahlmethode (manuell oder automatisiert) lassen sich in vier Kategorien einteilen. Die erste Kategorie umfasst die Menge der Tests, die zur Absicherung der Änderung durchgeführt werden und auch notwendig sind (richtig positiv, r_p). Die zweite Kategorie enthält die Menge der Tests, die zur Absicherung der Änderung durchgeführt werden, aber nicht notwendig sind (falsch positiv, f_p). Dies sind die Testfälle, die unnötig durchgeführt werden. Hier liegt damit theoretisches Einsparpotenzial vor. Die dritte Kategorie fasst die Tests zusammen, die nicht durchgeführt werden und auch obsolet sind (richtig negativ, r_n). Hierbei handelt es sich um die Tests, die die Auswahlmethode richtig ausgeschlossen hat. Die letzte Kategorie enthält die Tests, die nicht durchgeführt werden, obwohl sie notwendig sind (falsch negativ, f_n). Sie kennzeichnen das Risiko, das nach der Absicherung noch vorhanden ist. Die vier Kategorien sind in Tabelle 4-1 zusammenfassend dargestellt.

Tabelle 4-1: Wahrheitsmatrix

	Notwendige Tests	Obsolete Tests
	$r_p + f_n = N' $	$f_p + r_n = T \setminus N' $
Durchgeführte Tests $r_p + f_p = T' $	Richtig positiv $r_p = T' \cap N' $	Falsch positiv $f_p = T' \setminus N' $
Nicht durchgeführte Tests $f_n + r_n = T \setminus T' $	Falsch negativ $f_n = N \setminus T' $	Richtig negativ $r_n = (T \setminus T') \setminus N' $

4.3 Prinzipielle Testauswahlmethoden

Es gibt zwei grundsätzliche Möglichkeiten, Testfälle für die Absicherung von Änderungen auszuwählen.

¹¹² Bei Funktionserweiterungen werden ggf. zusätzliche Testfälle benötigt.

4.3.1 Inklusionsbasierte Testauswahl

Die inklusionsbasierte Testauswahlmethode stellt die Frage: Welche Testfälle müssen durchgeführt werden?

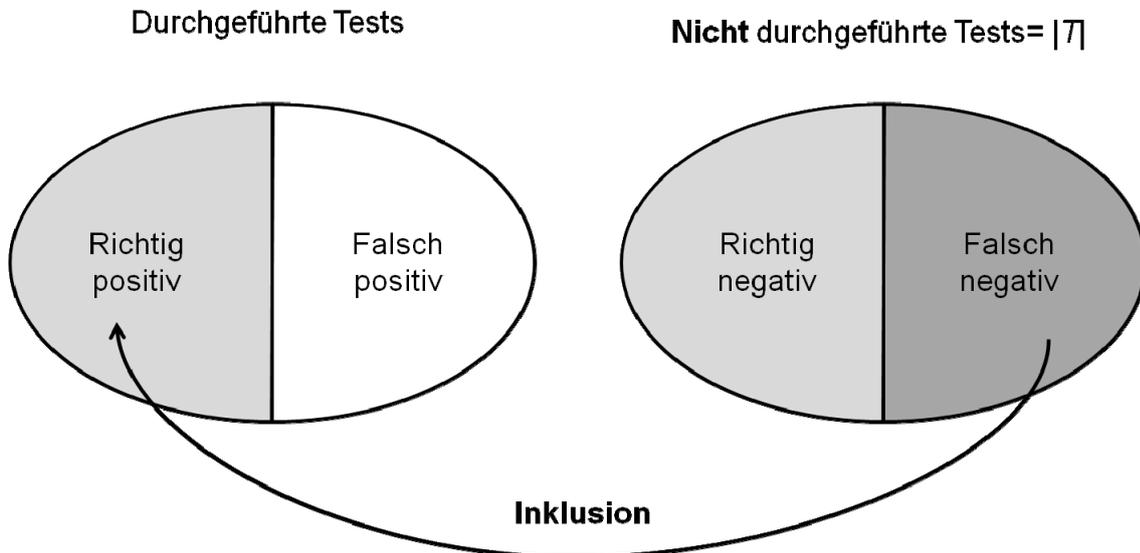


Abbildung 4-2: Prinzip der inklusionsbasierten Testauswahl

Das Ziel dieses Ansatzes ist daher eine Identifikation der Richtig-Positiven (r_p). Die Menge der ausgewählten Tests zur Überprüfung der Änderung ist zunächst leer, d.h. die Ausgangsbasis ist, dass nicht getestet wird:

$$T'_{\text{initial}} = \{ \} \rightarrow r_n + f_n = |T| \text{ bzw. } r_p + f_p = 0.$$

Auf Basis einer Auswirkungsanalyse, die die Effekte der Änderung bestimmt, werden die notwendigen Testfälle (r_p) aus T identifiziert und in die Menge der durchzuführenden Tests T' überführt. Dazu müssen möglichst viele falsch negative Tests (f_n) in richtig positive Tests (r_p) transferiert werden (siehe Abbildung 4-2).

4.3.2 Exklusionsbasierte Testauswahl

Die exklusionsbasierte Testauswahlmethode stellt die Frage: Welche Testfälle müssen **nicht** durchgeführt werden?

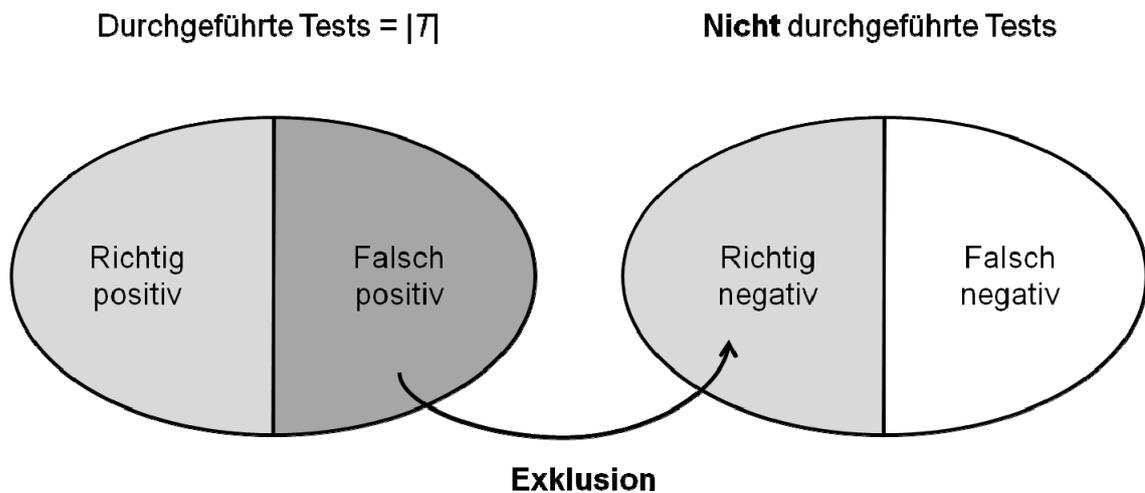


Abbildung 4-3: Prinzip der exklusionsbasierten Testauswahl

Das Ziel ist daher eine Identifikation der Richtig-Negativen (r_n). Bei diesem Ansatz entspricht die Menge der initial ausgewählten Tests zur Überprüfung einer Änderung der Gesamtmenge der Tests T , d.h. die Ausgangsbasis ist, dass alles getestet wird:

$$T'_{\text{initial}} = T \rightarrow r_p + f_p = |T| \text{ bzw. } r_n + f_n = 0.$$

Dazu müssen möglichst viele Falsch-Positive (f_p) in Richtig-Negative (r_n) überführt werden (siehe Abbildung 4-3).

4.4 Bewertung prinzipieller Testauswahlmethoden

Die beiden vorgestellten Ansätze besitzen unterschiedliche Eigenschaften. Die Auswirkungen dieser verschiedenen Charakteristiken auf das Ergebnis der Testauswahl werden im folgenden Abschnitt dargestellt. Dabei wird zunächst die Selektionsqualität untersucht, die anhand der Richtig-positiv-Rate sowie der Falsch-positiv-Rate analysiert wird. Im nächsten Schritt werden die Auswirkungen auf die Selektionskosten bestimmt sowie der Selektionsaufwand der beiden prinzipiellen Lösungsmöglichkeiten diskutiert. Als letztes Kriterium wird der Umgang mit mehreren Änderungen betrachtet.

4.4.1 Selektionsqualität

Die Selektionsqualität einer Testauswahlmethode kann anhand der Parameter Richtig-positiv-Rate und Falsch-positiv-Rate untersucht werden. Die Richtig-positiv-Rate (siehe Formel (4-1)) beschreibt den Anteil der bei der Testauswahl ermittelten notwendigen Tests an der Gesamtheit der notwendigen Tests. Sie stellt die Vollständigkeit der Testauswahl dar und kann damit als Maß für die Sicherheit der Methode herangezogen werden.

$$P_{\text{rp}} = \frac{r_p}{r_p + f_n} \quad (4-1)$$

Die Falsch-positiv-Rate (siehe Formel (4-2)) bezeichnet den Anteil der ermittelten nicht notwendigen Tests an der Gesamtmenge aller nicht notwendigen Tests. Sie gibt an, wie gut nicht notwendige Tests durch die Auswahl vermieden werden und ist damit ein Maß für die Fehlerrate der Methode.

$$P_{\text{fp}} = \frac{f_p}{r_n + f_p} \quad (4-2)$$

Die Receiver Operating Characteristic (ROC) bietet einen visuellen Ansatz für die Bewertung der Selektionsqualität. Die Methode kommt aus der Signalentdeckungstheorie, wo sie dazu eingesetzt wird, Signale von Rauschen zu unterscheiden. Die Technik wird heute in verschiedenen Anwendungsgebieten eingesetzt, u.a. Psychologie, Radiologie, Finanzwesen, Sozialkunde und maschinellem Lernen. ROC-Darstellungen bestehen aus einem zweidimensionalen Diagramm, bei dem die Richtig-positiv-Rate als Ordinate und die Falsch-positiv-Rate als Abszisse aufgetragen werden. Diese Illustration ermöglicht die Ermittlung des Zielkonfliktes zwischen der erfolgreichen Detektion von positiven Beispielen und der Missklassifizierung von negativen Beispielen. Für weitere Details zum ROC-Diagramm wird auf Tan¹¹³ verwiesen.

Abbildung 4-4 zeigt die ROC-Verläufe der beiden Testauswahlmethoden bei einer angenommenen, idealen Selektionsqualität (SQ) von 1 (siehe Formel (4-3)).

¹¹³ Tan: Receiver Operating Characteristic, 2009, S. 2349 ff.

$$SQ = \frac{r_p + r_n}{|T|} \quad (4-3)$$

Inklusionsbasierte Ansätze identifizieren falsch negative Beispiele aus der Menge der nicht ausgeführten Tests und überführen diese in richtig positive Tests. Die Vollständigkeit P_{rp} dieser Auswahlstrategie ist anfangs null, da noch keine Richtig-Positiven (r_p) in der Testmenge T' enthalten sind. Eine hohe Vollständigkeit der Testmenge T' kann mit der inklusionsbasierten Testmethode erreicht werden, wenn alle Falsch-Negativen (f_n) identifiziert, in Richtig-Positive (r_p) überführt und damit in die Testmenge T' integriert werden. Da die inklusionsbasierte Testauswahlmethode nur notwendige Tests selektiert, verbleiben alle obsoleten Tests in der Menge der nicht durchgeführten Tests ($T-T'$). Falsch positive Ergebnisse (f_p) werden so ausgeschlossen, d.h. $f_p = 0$. Die inklusionsbasierte Testauswahl ist damit immer exakt, d.h. die Fehlerrate ist immer null ($P_{fp} = 0$).

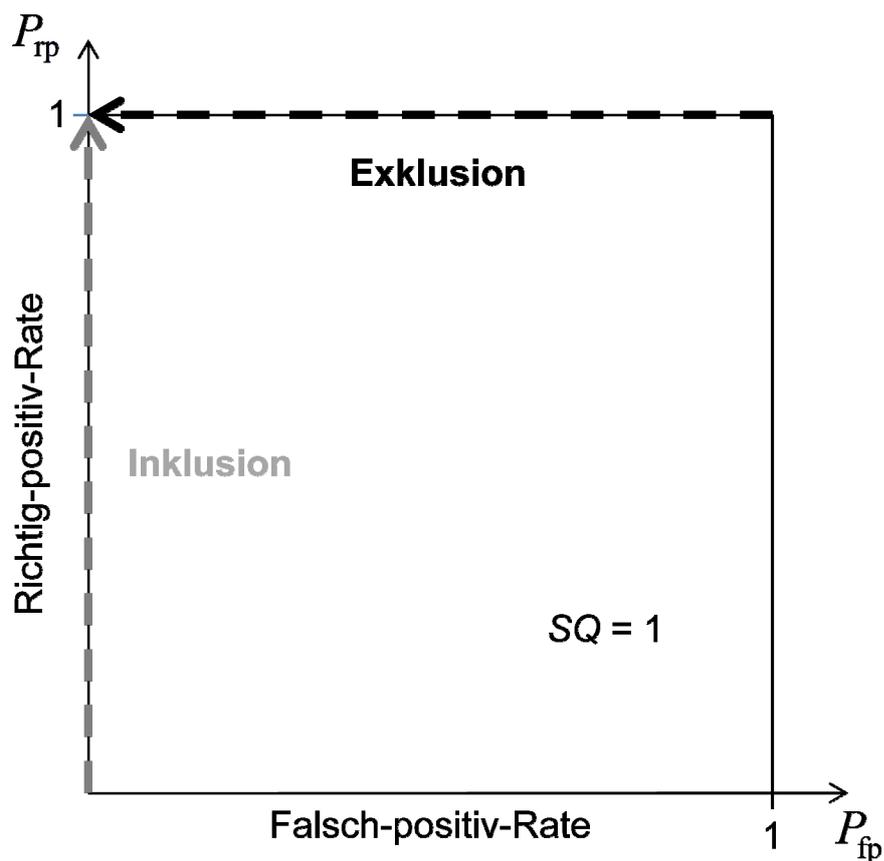


Abbildung 4-4: Vergleich der Selektionsqualität im ROC-Diagramm

Die exklusionsbasierte Testauswahl identifiziert falsch positive Beispiele aus der Menge der auszuführenden Tests und transferiert sie in Richtig-Negative. Die Strategie ist immer sicher i.S.v. Vollständigkeit. Alle notwendigen Tests zur Absicherung der Änderung bleiben in der Menge der durchgeführten Tests T' . Falsch negative Ergebnisse (f_n) werden so vermieden, d.h. $f_n = 0$. Die Fehlerrate dieser Auswahlstrategie beträgt anfangs 1, da noch keine Richtig-Negativen (r_n) in der Menge der nicht durchzuführenden Tests ($T-T'$) enthalten sind. Eine Fehlerrate von null kann erzielt werden, wenn alle falsch positiven Tests (f_p) bestimmt und in Richtig-Negative (r_n) transferiert werden.

Die ROC-Verläufe zeigen, dass beide Testauswahlmethoden den gleichen Punkt im ROC-Diagramm erreichen können ($P_{rp} = 1, P_{fp} = 0$), wenn ihre Selektionsqualität ideal ist. Die Startpunkte für die Analyse unterscheiden sich. Während der inklusionsbasierte Ansatz beim Punkt ($P_{rp} = 0, P_{fp} = 0$) beginnt, entspringt die ROC-Kurve für das exklusionsbasierte Konzept im Punkt ($P_{rp} = 1, P_{fp} = 1$).

4.4.2 Selektionsergebniskosten

Die Güte einer Testauswahlmethode kann auch anhand der zu erwartenden Selektionsergebniskosten evaluiert werden. Sie umfassen alle Kosten, die durch das Ergebnis der Testauswahl entstehen. Drummond und Holte¹¹⁴ schlagen ein Kostenmodell für Klassifikatoren im Bereich des maschinellen Lernens vor, das auf die Problemstellung dieser Arbeit übertragen werden kann. In Anlehnung an Drummond und Holte wird daher angenommen, dass die Kosten einer Testauswahl endlich und immer größer als Null sind. Ferner liegt dem hier angewandten Kostenmodell die Annahme zugrunde, dass die Kosten für die korrekte Einordnung eines Tests jederzeit geringer sind als die Kosten für die falsche Einordnung dieses Tests. Die bestmögliche Auswahltechnik klassifiziert jeden Test korrekt und besitzt erwartete Selektionsergebniskosten von null. Die zu erwartenden Selektionsergebniskosten (*SEK*) berechnen sich damit gemäß Formel (4-4).

$$SEK = P_{fn} \cdot P_+ \cdot C_{fn} + P_{fp} \cdot P_- \cdot C_{fp} \quad (4-4)$$

P_{fn} repräsentiert darin die Falsch-negativ-Rate ($f_n/(r_p+f_n)$), P_{fp} die Falsch-positiv-Rate ($f_p/(f_p+r_n)$). P_+ ist definiert als die Eintrittswahrscheinlichkeit, dass ein Test positiv ist, d.h. zur Menge der auszuführenden Tests gehört und $P_- = 1 - P_+$ bezeichnet die Eintrittswahrscheinlichkeit, dass ein Test negativ ist, d.h. zur Menge der nicht auszuführenden

¹¹⁴ Drummond, Holte: Explicitly representing expected cost, 2000

den Tests zählt.¹¹⁵ C_{fn} und C_{fp} erlauben eine Gewichtung der Falschklassifizierungskosten. C_{fn} charakterisiert dabei die Gewichtung der Kosten für eine falsche Einordnung von positiven Tests (f_n), während C_{fp} die Gewichtung der Kosten für eine nicht korrekte Zuordnung von negativen Tests (f_p) beschreibt.

Die maximalen erwarteten Kosten (SEK_{max}) treten auf, wenn alle Tests falsch klassifiziert werden, d.h. wenn $P_{fp} = 1$ und $P_{fn} = 1$ sind (siehe Formel ((4-5)).

$$SEK_{max} = P_+ \cdot C_{fn} + P_- \cdot C_{fp} \quad (4-5)$$

Die normalisierten erwarteten Kosten ($NSEK$) werden dann durch Division der Selektionsergebniskosten (SEK , Formel ((4-4)) durch die maximal möglichen erwarteten Kosten (SEK_{max} , Formel ((4-5)) bestimmt:

$$NSEK = \frac{P_{fn} \cdot P_+ \cdot C_{fn} + P_{fp} \cdot P_- \cdot C_{fp}}{P_+ \cdot C_{fn} + P_- \cdot C_{fp}} \quad (4-6)$$

Im ersten Schritt erfolgt keine Betrachtung unterschiedlicher Falschklassifizierungskosten. Unter dieser Randbedingung kann die Formel (4-6) folgendermaßen vereinfacht werden.

$$NESK_{simp} = P_{fn} \cdot P_+ + P_{fp} \cdot P_- \quad (4-7)$$

Die erwarteten Kosten der inklusionsbasierten Testauswahl beinhalten dann lediglich die Kostenanteile, die durch Falsch-Negative (f_n) entstehen, da Falsch-Positive (f_p) bei dieser Strategie vermieden werden. Da P_{fp} also immer null ist, muss daher nur der erste Kostensummand betrachtet werden. Die Kosten der exklusionsbasierten Auswahlstrategie bestehen nur aus den Kostenanteilen, die durch Falsch-Positive (f_p) verursacht werden, da Falsch-Negative (f_n) bei der Exklusion ausgeschlossen sind. Da P_{fn} also immer null ist, muss daher nur der zweite Kostensummand herangezogen werden. Wenn nun von einer idealen Selektionsqualität ausgegangen wird, sind am Ende des Testauswahlprozesses alle Beispiele korrekt eingeordnet. Somit entstehen weder bei dem einen noch

¹¹⁵ Die Wahrscheinlichkeiten $P(-)$ und $P(+)$ sind unbekannt. Ihre Höhe ist abhängig von der Auswirkungsgröße der zu untersuchenden Änderung. Da die Auswirkungsgröße sich stark unterscheiden kann, ist die zu entwickelnde Methode so auszulegen, dass sie für alle Verteilung von P_- zu P_+ anwendbar ist.

bei dem anderen prinzipiellen Lösungskonzept Selektionsergebniskosten und die Methoden sind als gleichwertig einzustufen.

Falls die Annahme einer idealen Selektionsqualität in der Praxis nicht zu realisieren ist, sind die Selektionsergebniskosten zunächst von den erzielbaren Fehlerraten P_{fn} (für die Inklusion) und P_{fp} (für die Exklusion) abhängig. Zusätzlich sind dann verschiedene Gewichtungen für C_{fn} und C_{fp} zu berücksichtigen. Wenn bei gleichen Fehlerraten die Kosten für falsch negative Tests C_{fn} größer sind als für falsch positive Tests C_{fp} , wird der Anwendungsbereich der exklusionsbasierten Testauswahlmethode vergrößert, da bei der Exklusion falsch negative Ergebnisse ganz vermieden werden und so nur weniger aufwändige, falsch positive Einstufungen anzurechnen sind. Falls die Kosten für falsch positive Tests C_{fp} höher ausfallen, ist dementsprechend die Inklusion als vorteilhaft zu bewerten.

In der Testauswahlsituation bei Änderungen auf Gesamtfahrzeugebene ist anzunehmen, dass sich die Höhe der Kosten unterscheidet. Falsch-Positive führen dazu, dass ein Test, der für die Absicherung der Änderung nicht notwendig ist, trotzdem durchgeführt wird. Die dadurch verursachten Kosten sind die Aufwände, die durch die unnötige Durchführung dieses Tests entstehen (Größenordnung 10^{3-4} €). Falsch-Negative führen zunächst dazu, dass ein Test, der für die Absicherung der Änderung notwendig ist, nicht durchgeführt wird. Dadurch müssen nicht zwangsläufig Probleme entstehen. Falls diese fehlende Testabdeckung jedoch bewirkt, dass ein Fehler im System nicht gefunden wird und es dadurch zu Funktionsbeeinträchtigung oder -schlimmer noch - zu Sicherheitsmängeln im Kundenbetrieb kommt, können die Kosten für diesen Fehler sehr hoch werden (im schlimmsten Fall Rückrufaktion mit Kosten in einer Größenordnung von 10^{6-8} €).

Da die beiden Methoden bisher unter der Annahme einer Selektionsqualität von 1 (d.h. fehlerfreie Testauswahl) untersucht wurden, spielt dieser Unterschied keine Rolle, da am Ende der Testauswahl bei beiden Strategien alle Tests richtig klassifiziert werden. Falls diese Annahme in der Praxis nicht realisierbar ist, sind unter dem Gesichtspunkt der Selektionsergebniskosten bei gleichen Fehlerraten falsch negative Einstufungen zwingend zu vermeiden. Um in der Gesamtfahrzeugabsicherung den Inklusionsansatz anwenden zu können, müsste dann nachgewiesen werden, dass das zugrundeliegende Modell zur Ermittlung der Auswirkungen der Änderung die betrachteten Wechselwirkungen vollständig und korrekt abbildet, so dass falsch negative Ergebnisse ausgeschlossen werden können.

4.4.3 Selektionsaufwand

Das Kostenmodell von Drummond und Holte¹¹⁶ berücksichtigt lediglich die Selektionskosten, die durch das Ergebnis der Testauswahl entstehen. Aufwände, die durch die Auswahl der Tests T' selbst verursacht werden, sind darin nicht enthalten und werden daher im Folgenden betrachtet.

Der Selektionsaufwand SA umfasst alle Aufwände, die durch die Selektion von Testfällen entstehen, z.B. Personalkosten, Rechenzeit oder benötigte Informationen. Er ist stark von der Umsetzung der Methode abhängig (z.B. Werkzeugunterstützung, Automatisierungsgrad). Um den Selektionsaufwand SA auf abstrakter Ebene bemessen zu können, ist eine vereinfachte Kenngröße erforderlich, die unabhängig von der Umsetzung ist.

Die Testauswahl teilt die Gesamtmenge der Tests T in zwei Teilmengen T' und $T - T'$ auf. Die Menge der durchzuführenden Tests T' enthält im Idealfall, der hier betrachtet wird, nur richtig positive Tests, während die Menge der nicht durchzuführenden Tests nur richtig negative Beispiele enthält. Der Selektionsaufwand setzt sich damit aus zwei Anteilen zusammen: dem Aufwand für die Identifizierung von richtig-positiven Tests sowie dem Aufwand für die Bestimmung von richtig negativen Tests (siehe Formel (4-8)).

$$SA_{\text{ges}} = SA_p(r_p) + SA_n(r_n) \quad (4-8)$$

Um die einzelnen Aufwandsbestandteile zu präzisieren, kann die Anzahl der Entscheidungen, die für die Integration oder den Ausschluss eines einzelnen Testfalls benötigt werden, herangezogen werden. Um einen Tests als notwendig bzw. nicht notwendig zu klassifizieren, muss geprüft werden, ob es eine, mehrere bzw. keine Wechselwirkungen dieses Tests mit der Änderung gibt. Die Gesamtanzahl der möglichen Wechselwirkungen k wird als endlich angenommen.¹¹⁷ Ferner wird zugrundegelegt, dass bei der Analyse jedes einzelnen Testfalls systematisch alle k Wechselwirkungen überprüft werden. Die Anzahl der Entscheidungen (Wechselwirkung mit der Änderung: ja oder nein) kann dann als Maß für den Selektionsaufwand SA herangezogen werden.

Inklusionsbasierte Ansätze identifizieren falsch negative Beispiele aus der Menge der nicht ausgeführten Tests und überführen diese in die Menge der richtig positiven Tests. Um einen Test als notwendig (r_p) zu identifizieren, werden Belege für eine Wechselwirkung des Tests mit der Änderung benötigt. Bei der Inklusion reicht es dazu aus, eine

¹¹⁶ Drummond, Holte: Explicitly representing expected cost, 2000

¹¹⁷ Vgl. Kap. 6.2.2

einzigste Wechselwirkung zu bejahen, um den Tests in die Menge der ausgeführten Tests T' zu überführen. Im besten Fall führt also die Überprüfung einer einzelnen Wechselwirkung schon zu der Aufnahme des Tests in die Menge der durchzuführenden Tests. Im aufwändigsten Fall wird die Wechselwirkung erst bei der Kontrolle der k -ten Wechselwirkung erkannt. Der Selektionsaufwand für die Identifizierung eines richtig positiven Tests $SA_p(r_p)$ kann daher gemäß Formel (4-9) zusammengefasst werden.

$$SA_p(r_p)_{\text{in}} \leq k \quad (4-9)$$

Um mit Hilfe der Inklusionsmethode einen Test als nicht notwendig (richtig negativ) einzustufen, müssen alle k Wechselwirkungen überprüft werden. Der Selektionsaufwand für die Bestimmung von richtig negativen Tests $SA_n(r_n)$ kann dann nach Formel (4-10) bestimmt werden.

$$SA_n(r_n)_{\text{in}} = k \quad (4-10)$$

Die exklusionsbasierte Testauswahl identifiziert falsch positive Beispiele aus der Menge der auszuführenden Tests und transferiert sie in Richtig-Positive. Um einen Test als nicht notwendig (r_n) einzuteilen, werden Beweise benötigt, dass dieser Test mit der Änderung in keiner Wechselbeziehung steht. Dazu sind alle k Wechselwirkungen zu überprüfen. Nur wenn alle Wechselwirkungen zu vernachlässigen sind, kann der Testfall in die Menge der nicht durchzuführenden Tests überführt werden. Der Selektionsaufwand für den Ausschluss von Testfällen mithilfe der Exklusionsmethode ist daher immer maximal hoch (siehe Formel (4-11)).

$$SA_n(r_n)_{\text{ex}} = k \quad (4-11)$$

Für die Integration eines Testfalls in die Menge der durchzuführenden Tests T' reicht es aus, eine Wechselwirkung zwischen der Änderung und dem Testfall zu finden. Der Selektionsaufwand für die Ermittlung von richtig positiven Tests kann daher gemäß Formel (4-12) erfasst werden.

$$SA_p(r_p)_{\text{ex}} \leq k \quad (4-12)$$

Damit ergibt sich für beide Methoden ein gleicher Gesamtselektionsaufwand SA_{ges} . Der gezeigte Ansatz zur Bestimmung des Selektionsaufwandes geht jedoch davon aus, dass die Aufwände für einen Beleg einer Wechselwirkung und für den Beweis, dass keine Wechselwirkung vorliegt, gleich hoch sind. Ist dies nicht der Fall, kann eine der beiden Methoden zu bevorzugen sein. Sollte z.B. der Beleg für eine Wechselwirkung einfacher zu realisieren sein, als der Beweis, dass keine Wechselwirkung existiert, ist der inklusionsbasierte Ansatz mit weniger Selektionsaufwand verbunden. Um dies zu überprüfen, ist allerdings eine detailliertere Betrachtungsebene notwendig.

4.4.4 Umgang mit mehreren Änderungen

Neben der Selektionsqualität und den Selektionskosten bzw. -aufwänden ist für die Bewertung einer Testauswahlmethode auch der Umgang mit mehreren Änderungen ein wichtiges Kriterium. Änderungen treten häufig nicht isoliert, sondern im Verbund auf.¹¹⁸ Es ist daher zu analysieren, ob und wie sich die beiden prinzipiellen Lösungskonzepte im Umgang mit Mehrfachänderungen unterscheiden. Die Anzahl der Änderungen wird mit n gekennzeichnet.

Umgang mit mehreren Änderungen bei der inklusionsbasierten Testauswahl

Bei der inklusionsbasierten Testauswahl werden aus der Gesamtmenge T die Tests identifiziert, die durchzuführen sind, d.h. es werden die richtig positiven Beispiele ermittelt. Für jede Änderung A_i wird also eine Menge von Tests $T_{\text{in},i}$ bestimmt, die durchgeführt werden. Dabei kann zu jeder Änderung A_i das Verhältnis b_i der Elemente von $T_{\text{in},i}$ zur Mächtigkeit der Gesamtmenge aller Tests T angegeben werden, das als Maß für die Auswirkungsgröße der Änderung herangezogen wird.

$$b_i = \frac{|T_{\text{in},i}|}{|T|} \quad (4-13)$$

Die Menge der Tests, die bei allen Änderungen (d.h. A_1 bis A_i) zusammen inkludiert werden, $T_{\text{in,ges}}$, kann als Vereinigung endlich vieler Mengen $T_{\text{in},i}$ aufgefasst werden.

$$T_{\text{in,ges}} = T_{\text{in},1} \cup T_{\text{in},2} \cup \dots \cup T_{\text{in},i} = \bigcup_{i=1}^n T_{\text{in},i} \quad (4-14)$$

Falls sich die Testbereiche einzelner Änderungen überschneiden, entsteht eine Schnittmenge. Sie repräsentiert die Tests, die eingespart werden können, wenn zwei Änderun-

¹¹⁸ Eckert et al.: Change and customization in complex engineering domains, 2003

gen gleichzeitig abgesichert werden. Die Vereinigung aller Schnittmengen, d.h. alle Tests, die in mindestens zwei $T_{in,i}$ vorkommen, wird als $T_{in,Überdeckung}$ bezeichnet.

$$T_{in,Überdeckung} = \bigcup_{\substack{i,j \\ i \neq j}} (T_{in,i} \cap T_{in,j}) \quad (4-15)$$

Der Überdeckungsgrad x charakterisiert das Verhältnis der Elemente von $T_{in,Überdeckung}$ zur Gesamtanzahl der Elemente von $T_{in,ges}$. Er kennzeichnet das Einsparpotenzial bei der parallelen Absicherung von mehreren Änderungen.

$$x = \frac{|T_{in,Überdeckung}|}{|T_{in,ges}|} \quad (4-16)$$

Die Menge der Tests, die bei allen Änderungen zusammen nicht durchgeführt wird, $T_{ex,ges}$, ergibt sich dann als Differenz von T und $T_{in,ges}$.

$$T_{ex,ges} = T \setminus T_{in,ges} \quad (4-17)$$

In Abbildung 4-5 sind beispielhafte Szenarien für den Umgang mit mehreren Änderungen dargestellt. Die zugrundeliegenden Parameter sind in Tabelle 4-2 zusammengefasst. Ausgangsbasis ist die Testmenge T mit einer Mächtigkeit von 100 Testfällen.

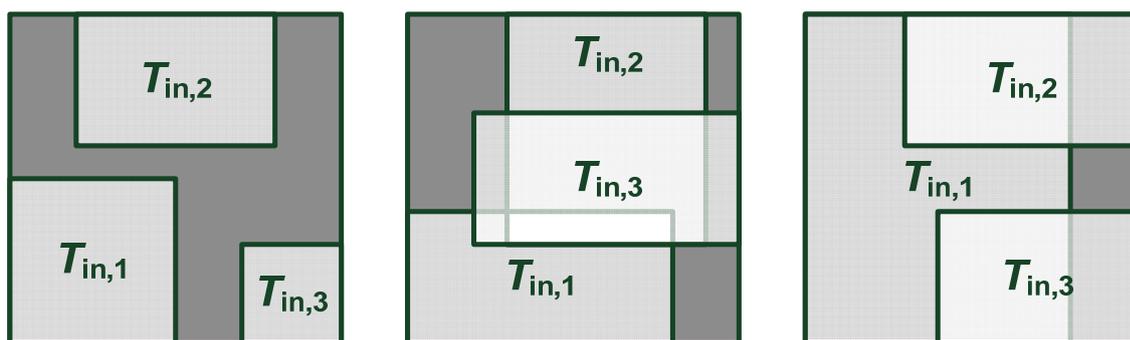


Abbildung 4-5: Beispielhafte Szenarien für mehrere Änderungen (Inklusion)

Das 1. Szenario zeigt drei Änderungen, deren Auswirkungen lokal begrenzt sind ($b_{max} = 0,25$). Die Besonderheit an diesem Szenario ist, dass die ermittelten Testmengen keine gemeinsame Schnittmenge besitzen, d.h. isoliert voneinander sind. Dennoch

müssen bei allen Änderungen zusammen nur etwas mehr als die Hälfte aller Test wiederholt werden, so dass sich die Testauswahl im Vergleich zur Vollabsicherung noch lohnt.

Tabelle 4-2: Beispielhafte Szenarien für mehrere Änderungen (Inklusion)

	<i>Szenario 1 (links)</i>	<i>Szenario 2 (Mitte)</i>	<i>Szenario 3 (rechts)</i>
$T_{in,1} / b_1$	25 / 0,25	32 / 0,32	80 / 0,8
$T_{in,2} / b_2$	24 / 0,24	42 / 0,42	28 / 0,28
$T_{in,3} / b_3$	9 / 0,09	32 / 0,32	24 / 0,24
$ T_{in,ges} $	58	76	96
x	0	$25/76 = 0,33$	$36/96 = 0,38$
$ T_{ex,ges} $	42	24	4

Das 2. Szenario repräsentiert wiederum drei Änderungen, die örtlich beschränkt sind und ca. 75 % aller Testfälle betreffen. Hier existiert jedoch eine Überdeckung der einzelnen Testmengen. In diesem Fall beträgt x ca. 33 %, was darauf hindeutet, dass das Einsparpotenzial bei der gemeinsamen Absicherung der drei Änderungen hoch ist. Szenario 3 zeigt eine Situation, in der eine Änderung (hier A_1) sehr große Auswirkungen hat. Zusammen mit den anderen Änderungen müssen 96 % aller Testfälle wiederholt werden. Hier ist das Einsparpotenzial im Vergleich zum Volltest als gering einzustufen. Es ist daher fraglich, ob sich in einer solchen Situation der Aufwand für die Testauswahl überhaupt noch lohnt oder ob es nicht weniger Aufwand ist, alle Tests durchzuführen.

Die Beispiele zeigen, dass die einzelnen Inklusionsmengen unabhängig voneinander sind. Ein hoher Überdeckungsgrad hilft zwar Testaufwand bei der parallelen Absicherung mehrerer Änderungen einzusparen (Szenario 2), ist jedoch nicht zwingend erforderlich (Szenario 1). Die Absicherung mehrerer Änderungen ist besonders vorteilhaft, wenn die Auswirkungen der einzelnen Änderungen lokal begrenzt sind und nur wenige Testfälle betreffen oder wenn ein hoher Überdeckungsgrad erzielt werden kann. Wenn einzelne Änderungen sehr große Auswirkungen haben, muss geprüft werden, ob sich der Aufwand für die Selektion im Vergleich zur Vollabsicherung noch lohnt (Szenario 3). Sobald die Änderungen zusammengenommen alle Testfälle in T betreffen, ist kein Einsparpotenzial im Vergleich zur Gesamtabsicherung mehr gegeben.

Umgang mit mehreren Änderungen bei der exklusionsbasierten Testauswahl

Bei der exklusionsbasierten Testauswahl werden aus der Gesamtmenge T die Tests identifiziert, die nicht durchzuführen sind, d.h. es werden die richtig negativen Beispiele ermittelt. Für jede Änderung A_i wird also eine Menge von Tests $T_{\text{ex},i}$ bestimmt, die nicht durchgeführt werden. Dabei kann zu jeder Änderung A_i das Verhältnis c_i der Elemente von $T_{\text{ex},i}$ zur Mächtigkeit der Gesamtmenge aller Tests T angegeben werden, das als Maß für die Begrenztheit der Auswirkungen der Änderung herangezogen wird.

$$c_i = \frac{|T_{\text{ex},i}|}{|T|} \quad (4-18)$$

Wenn mehrere Änderungen zeitgleich abgesichert werden, entfallen nur die Testfälle, die in allen $T_{\text{ex},i}$ enthalten sind. Daraus folgt, dass die Menge der Tests, $T_{\text{ex,ges}}$, die bei allen Änderungen (d.h. A_1 bis A_i) zusammen exkludiert werden, der Schnittmenge der einzelnen Exklusionsmengen entspricht.

$$T_{\text{ex,ges}} = T_{\text{ex},1} \cap T_{\text{ex},2} \cap \dots \cap T_{\text{ex},i} = \bigcap_{i=1}^n T_{\text{ex},i} \quad (4-19)$$

Die Menge der Tests, die bei allen Änderungen zusammen noch durchgeführt wird, $T_{\text{in,ges}}$, ergibt sich dann als Differenz von T und $T_{\text{ex,ges}}$.

$$T_{\text{in,ges}} = T \setminus T_{\text{ex,ges}} \quad (4-20)$$

In Abbildung 4-6 sind beispielhafte Szenarien für den Umgang mit mehreren Änderungen dargestellt. Die zugrundeliegenden Parameter sind in Tabelle 4-3 zusammengefasst. Ausgangsbasis ist wie zuvor eine Testmenge T mit einer Mächtigkeit von 100 Testfällen.

Das 1. Szenario zeigt drei Änderungen, deren Exklusionsmengen lokal begrenzt sind ($c_{\text{max}} = 0,25$). Die Besonderheit an diesem Szenario ist, dass die ermittelten Testmengen disjunkt sind, d.h. keine gemeinsame Schnittmenge besitzen. Damit können bei der parallelen Absicherung aller Änderungen keine Tests eingespart werden und die gesamte Testsuite muss erneut überprüft werden. Das 2. Szenario repräsentiert wiederum drei Änderungen, deren Exklusionsmengen örtlich beschränkt sind und ca. 75 % aller Testfälle betreffen. Die insgesamt große Exklusionsmenge kann jedoch nur teilweise ausgeschöpft werden, da die Schnittmenge der einzelnen Testmengen insgesamt nur 5 Testfälle umfasst, die bei der Absicherung im Vergleich zum Volltest entfallen können.

Szenario 3 zeigt eine Situation in der eine Änderung (hier A_1) sehr geringe Auswirkungen hat. Dennoch lassen sich bei der parallelen Absicherung aller drei Änderungen keine Testfälle einsparen, weil die Exklusionsmengen von A_2 und A_3 disjunkt sind.

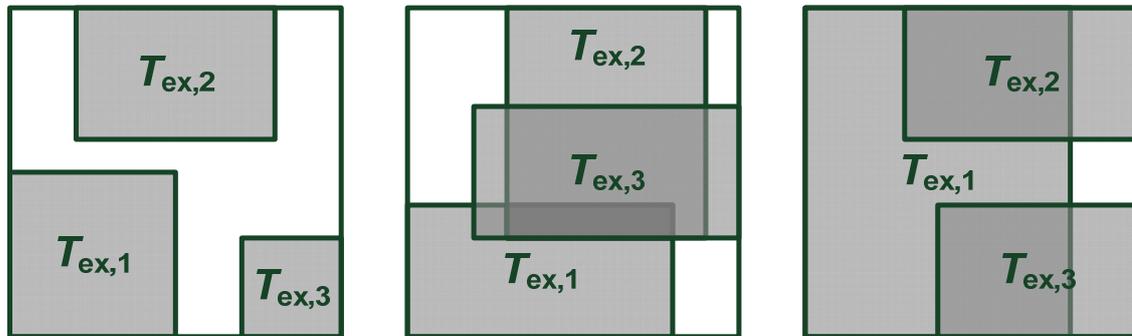


Abbildung 4-6: Beispielhafte Szenarien für mehrere Änderungen (Exklusion)

Tabelle 4-3: Beispielhafte Szenarien für mehrere Änderungen (Exklusion)

	<i>Szenario 1 (links)</i>	<i>Szenario 2 (Mitte)</i>	<i>Szenario 3 (rechts)</i>
$T_{ex,1} / c_1$	25 / 0,25	32 / 0,32	80 / 0,8
$T_{ex,2} / c_2$	24 / 0,24	42 / 0,42	28 / 0,28
$T_{ex,3} / c_3$	9 / 0,09	32 / 0,32	24 / 0,24
$ T_{ex,ges} $	0	5	0
$ T_{in,ges} $	100	95	100

Die Beispiele zeigen, dass bei der parallelen Absicherung von mehreren Änderungen mit Hilfe der exklusionsbasierten Testauswahl nur Testfälle eingespart werden können, wenn eine Schnittmenge zwischen den einzelnen Exklusionsmengen existiert. Die Mächtigkeit des Gesamtausschlusses richtet sich dabei nach der Größe der gemeinsamen Schnittmenge. Dies kann in der Praxis z.B. der Fall sein, wenn Änderungen ähnliche Module, Bauteile oder Funktionen betreffen. Wenn Änderungen sich in ihrer Art stark unterscheiden, ist die Wahrscheinlichkeit für eine gemeinsame Schnittmenge allerdings gering einzustufen. Daraus folgt für die Absicherung von mehreren Änderungen, dass die Exklusionstestfälle vor allem in den Bereichen der Schnittmengen zu suchen sind. Sobald der Nachweis erbracht werden kann, dass zwei Änderungen keine Schnittmenge haben, kann die exklusionsbasierte Testauswahl beendet werden. Dies

bedeutet, dass bei der Exklusion die Aussage darüber, dass alle Testfälle wiederholt werden müssen, ggf. früher erreicht werden kann.

Bei idealer Betrachtung der beiden Methoden wirken sich die Unterschiede nicht aus, da die Inklusionsmenge und die Exklusionsmenge immer komplementär sind. Das Ergebnis ist dann nur von der Zusammensetzung der einzelnen Änderungen abhängig (Auswirkungsgröße, Auswirkungsbereich usw.).

4.5 Zusammenfassung und Fazit

Für Testauswahlmethoden existieren zwei prinzipielle Lösungen. Bei der inklusionsbasierten Testauswahl werden die richtig positiven Tests (r_p) identifiziert und in die Menge der durchzuführenden Tests T' überführt. Die exklusionsbasierte Testauswahl hingegen bestimmt die richtig negativen Tests (r_n). Tabelle 4-4 fasst die unterschiedlichen Eigenschaften der beiden Konzepte zusammen.

Tabelle 4-4: Eigenschaften von inklusions- und exklusionsbasierter Testauswahl

	Inklusionsbasierte Strategie	Exklusionsbasierte Strategie
Vollständigkeit	Die Vollständigkeit dieser Auswahlstrategie ist anfangs null, da noch keine Richtig-Positiven (r_p) in der Testmenge enthalten sind. Eine hohe Vollständigkeit kann nur erreicht werden, wenn alle Falsch-Negativen (f_n) identifiziert und in Richtig-Positive (r_p) überführt werden.	Die exklusionsbasierte Testauswahl ist immer sicher i.S.v. Vollständigkeit. Alle notwendigen Tests zur Absicherung der Änderung bleiben in der Menge der durchgeführten Tests T' . Falsch negative Ergebnisse (f_n) werden so vermieden, d.h. $f_n = 0$.
Fehlerrate	Die inklusionsbasierte Testauswahl ist immer exakt. Alle obsoleten Tests verbleiben in der Menge der nicht durchgeführten Tests ($T-T'$). Falsch positive Ergebnisse (f_p) werden so ausgeschlossen, d.h. $f_p = 0$.	Die Fehlerrate dieser Auswahlstrategie beträgt anfangs 1, da noch keine Richtig-Negativen (r_n) in der Menge der nicht durchzuführenden Tests ($T-T'$) enthalten sind. Eine Fehlerrate von null kann nur erzielt werden, wenn alle Falsch-Positiven (f_p) bestimmt und in Richtig-

		Negative (r_n) transferiert werden.
Selektionskosten	Die Kosten der inklusionsbasierten Testauswahl beinhalten lediglich die Kostenanteile, die durch Falsch-Negative (f_n) entstehen, da Falsch-Positive (f_p) bei dieser Strategie vermieden werden.	Die Kosten der exklusionsbasierten Auswahlstrategie bestehen nur aus den Kostenanteilen, die durch Falsch-Positive (f_p) verursacht werden, weil Falsch-Negative (f_n) bei der Exklusion ausgeschlossen sind.
Selektionsaufwand	Für die Inklusion eines Testfalls wird mindestens ein Beweis für eine Wechselwirkung der Änderung mit dem Testfall benötigt.	Für die Exklusion eines Testfalls müssen alle k Wechselwirkungen der Änderung mit dem Testfall belegbar negiert werden.
Umgang mit mehreren Änderungen	Bei mehreren Änderungen kann die Vereinigungsmenge aller Inklusionsmengen $T_{in,i}$ als T' herangezogen werden. Die einzelnen Inklusionsmengen sind dabei unabhängig voneinander. Ein hoher Überdeckungsgrad hilft bei der parallelen Absicherung Aufwand im Vergleich zur mehreren Einzelabsicherungen einzusparen, ist jedoch nicht notwendige Voraussetzung.	Bei mehreren Änderungen bestimmt sich die Gesamtexklusionsmenge ($T - T'$) aus der Schnittmenge aller Exklusionsmengen. Die Größe der Ausschlussmenge richtet sich dabei nach der Mächtigkeit der Schnittmenge. Die einzelnen Änderungen sind daher nicht unabhängig voneinander. Schon wenn zwei Änderungen keine gemeinsame Schnittmenge besitzen, ist ein Volltest durchzuführen.

Konzepte, die auf der inklusionsbasierten Auswahlstrategie aufbauen, sind einerseits immer exakt, d.h. ihre Fehlerrate (Falsch-positiv-Rate) ist null. Andererseits sind sie nicht notwendigerweise vollständig. Die exklusionsbasierte Testauswahl produziert hingegen nur sichere Ergebnisse (i.S.v. Vollständigkeit), ist aber nicht automatisch frei von falsch positiven Resultaten. Die Kosten der beiden Strategien unterscheiden sich demzufolge nach ihrer Kostenart. Während bei der Inklusion Kosten für Falsch-Negative entstehen, verursacht die Exklusion nur Kosten für Falsch-Positive. Da die bisherigen Untersuchungen allerdings unter der Annahme einer perfekten Selektionsqualität von 1 (d.h. fehlerfreie Testauswahl) stattfanden, spielt dieser Unterschied keine

Rolle, weil am Ende der Testauswahl immer alle Tests richtig klassifiziert werden und so überhaupt keine Ergebniskosten entstehen. Es ist daher zu untersuchen, ob diese Annahme bei einer Konkretisierung der Ansätze zu halten ist bzw. welche Selektionsqualitäten mit beiden Konzepten zu erzielen sind. Mithilfe dieser Ergebnisse kann dann ein detaillierterer Vergleich der Selektionskosten angestellt werden.

Der Selektionsaufwand ist für beide Testauswahlkonzepte gleich hoch, solange der Aufwand für einen Beleg einer Wechselwirkung und der Aufwand für einen Beweis, dass keine Wechselwirkung vorliegt, gleich hoch sind. Es ist daher ferner zu untersuchen, ob diese Voraussetzung bei einer Detaillierung der Testauswahlmethoden noch gegeben ist bzw. welche Aufwände bei der Beweisführung entstehen. Damit kann dann eine genauere Bewertung des Selektionsaufwandes erfolgen.

Im Umgang mit mehreren Änderungen unterscheiden sich die beiden Methoden. Beim Inklusionsansatz sind die einzelnen Inklusionsmengen unabhängig voneinander. Die Gesamteinschlussmenge von Tests ergibt sich aus der Vereinigung aller individuellen Inklusionsmengen. Bei der exklusionsbasierten Testauswahl ist zwingend eine Schnittmenge aus allen Exklusionsmengen erforderlich, um Tests im Vergleich zum Vollabsicherung einzusparen. Die einzelnen Änderungen sind damit nicht unabhängig voneinander. Die Unterschiede wirken sich bei idealer Betrachtung der beiden Methoden nicht aus, da Inklusionsmenge und die Exklusionsmenge immer komplementär sind. Das Ergebnis der Testauswahl ist dann lediglich von der Zusammensetzung der einzelnen Änderungen abhängig (Auswirkungsgröße, Wirkungsbereich usw.). Wie sich die verschiedenen Eigenschaften bei praktisch erzielbaren Selektionsqualitäten auswirken, ist allerdings noch zu untersuchen.

5 Inklusionsanalyse

Die Ausführungen im vorangegangenen Kapitel haben gezeigt, dass für einen detaillierteren Vergleich eine Konkretisierung der beiden prinzipiellen Lösungskonzepte für die Testauswahl zur Absicherung von Änderungen auf Gesamtfahrzeugebene erforderlich ist. In diesem Kapitel wird daher das Konzept der Inklusion weiter ausgearbeitet.

5.1 Prinzip der Inklusionsbeweissführung

Die Ausgangsbasis für die Inklusion bildet die initiale Testmenge T , die alle Tests enthält, die zur Absicherung des Systems durchgeführt werden. Ein Test t_i kann in die Menge der durchzuführenden Tests T' überführt werden, wenn der Nachweis erbracht wird, dass der Test zur Absicherung der Änderung notwendig ist. Notwendig sind prinzipiell alle Tests aus T , die Fehler identifizieren, die durch die Änderung A in das Gesamtsystem eingebracht wurden, sog. fehleraufdeckende („fault-revealing“) Tests. Für den erforderlichen Nachweis wird das Ergebnis $R(t_i)$ des Tests t_i betrachtet.

Das Ergebnis $R(t_i)$ evaluiert den Ausgang des Tests t_i , d.h. das Verhalten des Testobjektes, das bei der Durchführung des Tests unter den spezifizierten Bedingungen produziert wird.¹¹⁹ Testergebnisse können auf drei verschiedenen Messskalen bestimmt werden.¹²⁰ Auf einer Nominalskala können Testergebnisse in verschiedene Kategorien klassifiziert werden, im einfachsten Fall z.B. in bestanden und nicht bestanden. Auf einer Ordinalskala werden die Testresultate in einer Rangfolge arrangiert. Ein Beispiel für diese Klasse ist die ATZ Skala¹²¹, die verwendet wird, um Fahrzeugeigenschaften (z.B. Agilität, Stabilität, Komfort) auf Stufen von 1 bis 10 zu bewerten. Die dritte Kategorie bildet die metrische Skala (auch Kardinalskala genannt). Beispiele für diese Messskala sind der Bremsweg oder der streckenbezogene Kraftstoffverbrauch.

Um nachzuweisen, dass ein Test notwendig ist, wird der Einfluss der Änderung A auf das Ergebnis R des Tests t_i analysiert. Die Auswirkung der Änderung auf das Testergebnis kann in unterschiedliche Gruppen eingeteilt werden. Die Ergebnisse des Tests können gleich sein, d.h. $R(t_i, A) = R(t_i)$; die Resultate können kleiner gleich (\leq) oder größer gleich (\geq) sein, d.h. $R(t_i, A) \leq R(t_i)$ oder $R(t_i, A) \geq R(t_i)$. Wobei ($<$) schlechtere

¹¹⁹ ISTQB Glossary, 2015

¹²⁰ Kohn: Statistik, 2005, S. 13-15

¹²¹ Aigner: Zur zuverlässigen Beurteilung von Fahrzeugen, 1982

Testergebnisse symbolisiert, während ($>$) für bessere Testausgänge steht. Darüber hinaus können die Testergebnisse auch definitiv ungleich sein, d.h. $R(t_i, A) \neq R(t_i)$.

In Anlehnung an diese Unterscheidung, muss ein Testfall in die Menge der durchzuführenden Tests integriert werden, wenn nachgewiesen werden kann, dass das Testergebnis inklusive der Änderung $R(t_i, A)$ schlechter als das Testergebnis $R(t_i)$ ohne die Änderung ist. Bei einem Testresultat auf der Nominalskala kann nur die Ungleichheit herangezogen werden.

Für die Inklusion kann daher folgender Einschlussbeweis eingesetzt werden.

Einschlussbeweis I: Ergebnisverschlechterung

Ein Test t_i ($t_i \in T$) muss in die Menge der durchgeführten Tests T' integriert werden, wenn ein formaler Beweis existiert, der demonstriert, dass die Änderung A einen negativen Einfluss auf das Ergebnis R_i des Tests t_i hat.

$$\text{EB I:} \quad R(t_i, A) \overset{!}{<} R(t_i) \rightarrow t_i \in T_{\text{in},1} \quad (5-1)$$

Eine Änderung A hat negative Auswirkungen auf einen Test t_i , wenn das Testergebnis $R(t_i, A)$ um mehr als der Vertrauensbereich des Testergebnisses $R(t_i)$ schlechter ist.

Um die Ergebnisverschlechterung nachzuweisen, wird ein formaler Beweis benötigt, der zeigt, dass der Test von der Änderung nachteilig betroffen ist. Nach Rothermel¹²² gibt es jedoch keine effektive Methode, um diese fehleraufdeckenden Testfälle zu ermitteln. Unter bestimmten Rahmenbedingungen¹²³ kann jedoch eine Übermenge von Testfällen gebildet werden, sog. modifikationsdurchlaufende („modification-revealing“) Tests. In dieser Übermenge sind dann alle fehleraufdeckenden, d.h. nachteilig betroffenen Tests enthalten. Gleichzeitig können jedoch auch falsch positive Testfälle, d.h. Testfälle, die keine Fehler aufdecken, nicht ausgeschlossen werden. Eine fehlerfreie Umsetzung der inklusionsbasierten Methoden kann daher nicht garantiert werden.

¹²² Rothermel: Efficient, Effective Regression Testing, 1996

¹²³ Annahme eines kontrollierten Regressionstests, d.h. alle Faktoren außer dem geänderten Quellcode, die den Output des Testfalls beeinflussen, bleiben konstant. Für Details wird verwiesen auf: Rothermel: Efficient, Effective Regression Testing, 1996

5.2 Stand der Technik

In der Softwaretechnik sind jene Testfälle von einer Änderung betroffen, die einen neuen oder modifizierten Quellcode betreffen oder die ehemals Quelltext ausgeführt haben, der im neuen System nicht mehr vorhanden ist.¹²⁴ In der Literatur sind viele systematische Ansätze für die inklusionsbasierte Auswahl von Testfällen auf Softwareebene beschrieben. Detaillierte Zusammenfassungen sind z.B. bei Engström et al.¹²⁵ und Yoo und Harman¹²⁶ zu finden. Die Techniken ermitteln die Testfälle auf Basis einer Änderungsanalyse („change impact analysis“)¹²⁷, die wiederum meist auf den Quelltext des Systems zurückgreift. Einige der beschriebenen Ansätze sind sicher („safe“)¹²⁸, d.h. sie wählen immer alle modifikationsdurchlaufenden Tests aus.¹²⁹ Rothermel und Harrold konnten zudem nachweisen, dass einige dieser sicheren Techniken sehr effizient einsetzbar sind, da sie automatisiert durchgeführt werden und auf Informationen zurückgreifen, die ohnehin im Entwicklungsprozess entstehen.¹³⁰ Dennoch werden diese Konzepte in der Industrie wenig angewandt.¹³¹ Dies liegt vor allem daran, dass Testpersonen häufig keinen Zugriff auf den Code haben und daher nur solche Methoden praktische Relevanz besitzen, die nicht auf dem Quelltext aufbauen (sog. Black-Box-Ansätze).¹³² Beispiele für diese Kategorie sind bei Orso et al.¹³³ und Sajeev und Wibowo¹³⁴ zu finden. Nachteil der beiden Konzepte ist allerdings, dass sie nicht sicher i. S. der Definition von Rothermel sind.¹³⁵

¹²⁴ Nörenberg: Effizienter Regressionstest von E/E-Systemen, 2012, S. 71

¹²⁵ Engström et al.: A systematic review on regression test selection techniques, 2010

¹²⁶ Yoo, Harman: Regression testing minimization, selection and prioritization, 2010

¹²⁷ Vgl. dazu auch Kap. 3.4.2

¹²⁸ Die Definition von „sicher“ („safe“) geht auf Rothermel zurück, vgl. Rothermel: Analyzing Regression Test Selection Techniques, 1996. Danach sind Testauswahlmethoden sicher, wenn sie immer alle modifikationsdurchlaufenden Tests identifizieren.

¹²⁹ Eine Analyse sicherer Testauswahlmethoden ist bei Rothermel (Rothermel: Empirical studies of safe regression test selection techniques, 1998) zu finden.

¹³⁰ Rothermel, Harrold: Empirical Studies of a Safe Regression Test Selection Technique, 1998

¹³¹ Vgl. z.B. Engström, Runeson: A qualitative survey of regression testing practices, 2010

¹³² Engström et al.: Improving Regression Testing Transparency and Efficiency, 2011

¹³³ Orso et al.: Using component metadata to regression test component-based software, 2007

¹³⁴ Sajeev, Wibowo: Regression test selection based on version changes of components, 2003

¹³⁵ Engström et al.: Empirical Evaluations of Regression Test Selection Techniques, 2008

Auf Fahrzeugesamtebene sind alle Tests zu ermitteln, die Systembestandteile betreffen, die in Wechselwirkung mit der Änderung A stehen. Auf dieser Betrachtungsstufe existiert kein Quelltext, auf den zur Analyse der Wechselbeziehungen der Änderung zurückgegriffen werden kann. Es wird daher ein Modell benötigt, anhand dessen die Wechselwirkungen der Änderung nachvollzogen werden können. Unter Wechselwirkungen werden dabei in Anlehnung an Schramm et al.¹³⁶ alle internen Funktionszusammenhänge und physikalischen Gesetze verstanden. Ziel der Modellbildung ist, eine mathematisch-analytische Beschreibung des Systems, in diesem Falle also des Gesamtfahrzeugs, zu erhalten, die es erlaubt, die Auswirkungen von Änderungen zu untersuchen. Bei heutigen Gesamtfahrzeugsystemen handelt es sich häufig um mechatronische Systeme, die auf einer funktions- und hardwareorientierten Kombination von mechanischen, elektrischen und elektronischen Komponenten und Teilsystemen sowie der entsprechenden Software beruhen.¹³⁷ Das zugrundeliegende Modell muss daher in der Lage sein, diese heterogenen Eigenschaften und Strukturen gemeinsam abzubilden.

Der Einsatz von komplexen mathematischen Fahrzeugmodellen hat in den letzten Jahren sowohl in der Vorentwicklung, der eigentlichen Entwicklung und der anschließenden Optimierung von Produkten stark an Bedeutung gewonnen.¹³⁸ Viele Systeme sind kommerziell verfügbar (z.B. folgende Systeme zur Simulation der Fahrzeugdynamik: CarMaker von IPB¹³⁹, veDYNA von TESIS Dynaware¹⁴⁰ oder CarSim von Mechanical Simulation¹⁴¹). Modelle können mit unterschiedlicher Komplexität und Aussagekraft gestaltet werden.¹⁴² Je komplexer das Modell aufgebaut ist, desto genauer können damit die Abhängigkeitsbeziehungen beschrieben werden. Die hohe Komplexität führt jedoch fast zwangsläufig dazu, dass aufwändigere und in der Regel nichtlineare Modellgleichungen sowie ein hoher Rechenzeitbedarf entstehen.¹⁴² Außerdem wächst mit zunehmender Modellkomplexität auch die Anzahl der benötigten Modellparameter.¹⁴² Typische Quellen für Parameter in der Fahrzeugtechnik sind nach Schramm et al.¹⁴² CAD-Modelle (z.B. für Massen, Abmessungen, Trägheitsmomente), direkte Messungen von Parametern (z.B. Massen, Trägheitsmomente, Feder- und Dämpferkennlinien), Annahmen, Schätzungen und ggf. Identifikationsverfahren (für schwer bestimmbar oder

¹³⁶ Schramm et al.: Modellbildung und Simulation der Dynamik von Kraftfahrzeugen, 2013, S. 5

¹³⁷ Schramm et al.: Modellbildung und Simulation der Dynamik von Kraftfahrzeugen, 2013, S. 1 f.

¹³⁸ Vgl. Schramm et al.: Modellbildung und Simulation der Dynamik von Kraftfahrzeugen, 2013, S. 1 f.; Westermann: Modellbildung und Simulation, Springer Verlag, Berlin Heidelberg, 2010, S. 1

¹³⁹ IPG: Car Maker (<http://ipg.de/simulationsolutions/carmaker/>)

¹⁴⁰ TESIS: veDYNA (<http://www.thesis-dynaware.com/produkte/vedyna/uebersicht.html>)

¹⁴¹ Mechanical Simulation: CarSim (<http://carsim.com/products/carsim/index.php>)

¹⁴² Schramm et al.: Modellbildung und Simulation der Dynamik von Kraftfahrzeugen, 2013, S. 4

unscharfe Größen wie z.B. Reibungseffekte und Elastizitäten) oder die Berechnung von Parametern und Kennfeldern durch andere Berechnungs- und Simulationsmethoden (z.B. Finite-Elemente-Methode (FEM)). Häufig übersteigt der Aufwand bei der Beschaffung der relevanten Parameter den Aufwand für die Modellerstellung deutlich.¹⁴³ Das Problem jeglicher Modelle ist, dass sie stets mit Idealisierungen und Abstraktionen verbunden sind und daher nie ein absolut vollständiges Abbild eines Systems darstellen.¹⁴⁴ In Tabelle 5-1 ist die Prognosegüte von CAE-Methoden im Bereich Gesamtfahrzeug dargestellt (A: im Einsatz, prognosesicher; B: im Einsatz, zur Unterstützung der Entwicklung; C: beschränkter Einsatz wegen unvollständiger Methode).

Tabelle 5-1: Prognosegüte von CAE-Methoden im Bereich Gesamtfahrzeug nach Meywerk¹⁴⁵

	<i>A</i>	<i>B</i>	<i>C</i>
Fahrdynamik		x	
Fahrleistung / Verbrauch	x		
Schwingungen, linear	x		
Schwingungen, nichtlinear		x	
Lebensdauer			x
Außenströmung		x	
Innenströmung	x		
Motorraumdurchströmung	x		
Klimatisierung			x
Wärmemanagement			x
Regelungstechnik		x	

Die Auflistung offenbart, dass zwar für viele Fahrzeugbereiche Modelle vorhanden sind, viele dieser Modelle jedoch nicht prognosesicher sind.

¹⁴³ Schramm et al.: Modellbildung und Simulation der Dynamik von Kraftfahrzeugen, 2013, S. 4

¹⁴⁴ Schramm et al.: Modellbildung und Simulation der Dynamik von Kraftfahrzeugen, 2013, S. 5; Roddeck, W.: Einführung in die Mechatronik. Vieweg+Teubner Verlag, Wiesbaden, 2012, S. 28

¹⁴⁵ Meywerk: CAE-Methoden in der Fahrzeugtechnik, 2007, S. 3

5.3 Fazit

Die vorangegangenen Betrachtungen zeigen, dass für die inklusionsbasierte Testauswahl nach dem Stand der Technik und Forschung keine effektive Methode existiert, mit der alle fehleraufdeckenden Tests ermittelt werden können. Es ist lediglich möglich, eine Übermenge von Testfällen zu bestimmen, die von der Änderung betroffen sind (sog. modifikationsdurchlaufende Tests). Dabei können jedoch falsch positive Testfälle nicht ausgeschlossen werden. Vorteil vieler in der Literatur beschriebener und in der Praxis eingesetzter Verfahren der inklusionsbasierten Testauswahl ist, dass sie sehr effizient einsetzbar sind. Die zugrundeliegenden Modelle für die Testauswahl sind teilweise schon vorhanden (z.B. Quelltext, CAD-Modelle etc.) und auch die Testauswahl ist automatisiert durchführbar, so dass der Selektionsaufwand als gering eingestuft werden kann. Nachteil der inklusionsbasierten Testauswahl ist jedoch, dass die Qualität der Testauswahl von der Qualität des zugrundeliegenden Modells abhängig ist. Bei Verfahren, die die Auswirkungen von Änderungen in der Software analysieren, kann teilweise eine sehr hohe Selektionsqualität erzielt werden (Verfahren, die z.B. auf dem Quelltext aufbauen).

Um in der Absicherung von Änderungen auf Gesamtfahrzeugebene die inklusionsbasierte Testauswahl einsetzen zu können, muss nachgewiesen werden, dass das zugrundeliegende Modell des Gesamtfahrzeugs zur Ermittlung der Wechselwirkungen ein vollständiges und korrektes Abbild des Gesamtsystems darstellt. Vollständig ist ein Modell in diesem Sinne, wenn es alle Wechselwirkungen der Änderung erfasst. Aufgrund der Gesamtsystemgröße und der Vielfalt und Komplexität der Wechselwirkungen der einzelnen Teilsysteme ist dies eine schwierige Aufgabe, da jede übersehene Wechselwirkung zu einer Reduktion der Testabdeckung führt. Korrekt ist das Modell darüber hinaus, wenn damit die Testergebnisveränderung in jedem Fall feststellbar ist. Das Modell muss also nicht nur die Art der Wechselwirkung, sondern auch die Größe der Auswirkung abbilden, um feststellen zu können, wie genau sich die Änderung auf das Testergebnis auswirkt. Die Anwendung der inklusionsbasierten Testauswahl zur Absicherung von Änderungen auf Gesamtfahrzeugebene ist daher nur unter der Voraussetzung möglich, dass ein Beweis für die Vollständigkeit und Korrektheit des zugrundeliegenden Modells vorliegt.

Da die Menge der Wechselwirkungen unbekannt ist und absolute Aussagen über eine unbestimmte oder unendlich große Menge gemäß Popper¹⁴⁶ grundsätzlich nicht möglich sind, kann dieser Beweis lediglich induktiv geführt werden, d.h. anhand von Beobach-

¹⁴⁶ Popper: Logik der Forschung, 2005

tungen oder Erfahrungen. Es ist daher offen, wie die inklusionsbasierte Testauswahl auf Gesamtfahrzeugebene jemals beweisen kann, dass sie diese Voraussetzungen erfüllt.

6 Exklusionsanalyse

Nachdem in Kapitel 5 die Inklusionsanalyse konkretisiert wurde, erfolgt in diesem Kapitel die weitere Ausgestaltung des Exklusionskonzeptes. Dazu wird zunächst das Prinzip der Exklusionsbeweissführung näher erläutert. Anschließend wird dann eine exemplarische Umsetzung des Ansatzes beschrieben.

6.1 Prinzip der Exklusionsbeweissführung

Eine Exklusion von Testfällen erfolgt durch einen Prozess der Elimination. Ausgangsbasis für den Ausschluss bildet die initiale Testmenge T , die alle Tests enthält, die zur Absicherung des Systems durchgeführt werden. Ein Test t_i kann von der Menge der durchgeführten Tests T ausgeschlossen und in die Menge der exkludierten Tests T_{ex} überführt werden, wenn der Nachweis erbracht wird, dass der Test zur Absicherung der Änderung nicht notwendig ist. Für diesen Nachweis wird wie in Abschnitt 5.1 das Ergebnis $R(t_i)$ des Tests t_i betrachtet.

Ein Testfall kann dann von der Menge der durchzuführenden Tests exkludiert werden, wenn nachgewiesen wird, dass das Testergebnis inklusive der Änderung $R(t_i, A)$ gleich oder besser als das Testergebnis $R(t_i)$ ohne die Änderung ist. Bei einem Testresultat auf der Nominalskala kann nur die Gleichheit für die Exklusion herangezogen werden. Wenn das Testergebnis schlechter ausfällt, kann der Test nur unter der Bedingung entfallen, dass ein anderer Test durchgeführt wird, der diese Verschlechterung ebenfalls aufzeigt.

Für die Exklusion können daher zwei Ausschlussbeweise herangezogen werden.

Ausschlussbeweis I: Ergebnisneutralität

Ein Test t_i ($t_i \in T$) kann von der Menge der durchgeführten Tests T ausgeschlossen werden, wenn ein formaler Beweis existiert, der demonstriert, dass die Änderung A keinen negativen Einfluss auf das Ergebnis R_i des Tests t_i hat.

$$\text{AB I:} \quad R(t_i, A) \stackrel{!}{\geq} R(t_i) \rightarrow t_i \in T_{\text{ex},1} \quad (6-1)$$

Eine Änderung A hat keine negative Auswirkung auf einen Test t_i , wenn das Testergebnis $R(t_i, A)$ innerhalb des Vertrauensbereichs des Testergebnisses $R(t_i)$ liegt oder besser

ist. Um die Ergebnisneutralität nachzuweisen, wird ein formaler Beweis benötigt, der zeigt, dass der Test von der Änderung nicht nachteilig betroffen ist. Wie dieser Nachweis aussehen kann, wird beispielhaft in Abschnitt 6.2.1 und Abschnitt 6.2.2 gezeigt.

Ausschlussbeweis II: Ergebnisgleichheit

Ein Test t_i ($t_i \in T$) kann von der Menge der durchgeführten Tests T ausgeschlossen werden, wenn eine Kombination anderer Tests $t_j \dots t_m$ ($t_j \dots t_m \in T$) existiert, deren Ergebnisse $\bar{R}_j \dots \bar{R}_m$ zusammengenommen den Einfluss der Änderung A mindestens genauso gut zeigen wie das Ergebnis \bar{R}_i des Tests t_i .

$$\text{AB II:} \quad \bar{R}(t_i, A) \leq \sum_{l=j}^m \bar{R}(t_l, A) \rightarrow t_i \in T_{\text{ex},2} \wedge t_j \dots t_m \in T_{\text{in},2} \quad (6-2)$$

Der Beweis der Ergebnisgleichheit markiert einen Test t_i als redundant, wenn die Ergebnisse einer Kombination anderer Tests aus der durchzuführenden Testmenge die Auswirkungen der Änderungen mindestens genauso gut zeigen wie das Ergebnis von t_i . Dies ist der Fall, wenn bei den anderen Tests mindestens für jede Dimension \bar{R}_i von t_i ein Test vorhanden ist, der die Auswirkungen der Änderung offenbart. Die Tests, die als Referenz für die Exklusion mithilfe dieses Ausschlussbeweises herangezogen werden, dürfen allerdings nicht weiter ausgeschlossen werden. Sie verbleiben in der Menge der durchzuführenden Tests und erfahren damit eine Inklusion. Eine beispielhafte Umsetzung des Ausschlussbeweises Ergebnisgleichheit ist in Abschnitt 6.2.3 dargestellt.

Das resultierende Gesamtkonzept für die exklusionsbasierte Testauswahlmethode zeigt Abbildung 6-1. Eingangsgrößen für die Auswahlmethode bilden die initiale Testmenge T und die Änderung A , die abgesichert werden soll. In einem ersten Schritt wird der erste Ausschlussbeweis (AB I: Ergebnisneutralität) evaluiert. Alle Testfälle, die nachweislich nicht von der Änderung betroffen sind, werden eliminiert und in die Menge der nicht durchgeführten Tests $T_{\text{ex},1}$ überführt. Die Testmenge wird so reduziert. Die verbleibenden Tests $T_{\text{red},1}$ werden in einem zweiten Schritt weiter analysiert. Grundlage für diesen Reduktionsschritt bildet der zweite Ausschlussbeweis (AB II: Ergebnisgleichheit). Alle Tests, deren Ergebnisse bzgl. der Änderung durch die in der Testmenge verbleibenden Tests $T_{\text{red},2}$ abgedeckt sind, werden ausgeschlossen und in die Menge der exkludierten Tests $T_{\text{ex},2}$ transferiert.

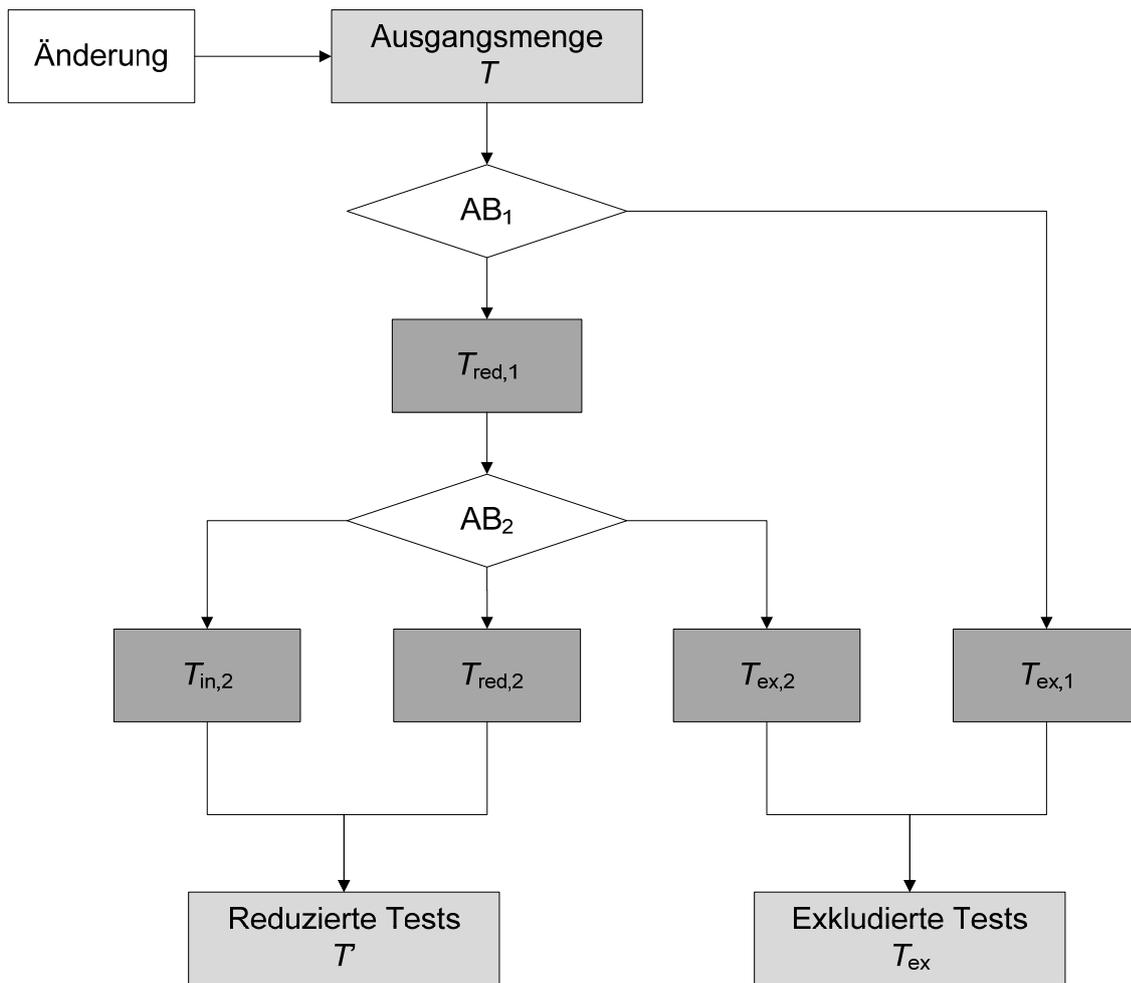


Abbildung 6-1: Konzept der exklusionsbasierten Testauswahlmethode

Durch die genannte Reihenfolge der Ausschlussbeweise wird ermöglicht, dass die initiale Testmenge zunächst verringert wird (AB I), bevor mit der verbleibenden, reduzierten Testmenge $T_{\text{red},1}$ eine Vergleichsanalyse durchgeführt wird (AB II). Die Vergleichsmenge ist so möglichst gering, was den Aufwand für die Bewertung verringert. Eine umgekehrte Reihenfolge ist darüber hinaus nicht zielführend, da die Testfälle, die als Referenz für den zweiten Ausschlussbeweis herangezogen wurden, nicht weiter exkludiert werden dürfen.

Zusammengenommen liegt damit ein deduktiv abgeleiteter Ansatz vor, der bei Änderungen eine sichere Identifikation von Testfällen aus einer Ausgangsmenge erlaubt. Wie eine Konkretisierung dieses Konzeptes für Absicherungen von Änderungen im Gesamtfahrzeug aussehen kann, zeigen die folgenden Abschnitte.

6.2 Exemplarische Umsetzung

In diesem Kapitel wird eine beispielhafte Umsetzung der exklusionsbasierten Testauswahlmethode beschrieben. Der dargestellte Ansatz stellt mögliche Ausgestaltungen der Ausschlussbeweismethode vor. Damit wird gezeigt, dass das Konzept auf Änderungsabsicherungen auf Gesamtfahrzeugebene übertragbar ist. Andere Realisierungen oder Erweiterungen sind möglich.

Der Stand der Technik bildet die Ausgangslage für die Exklusion. Die Testmenge T enthält daher alle Tests, die vom OEM und den Zulieferern im Absicherungsprozess auf Gesamtfahrzeugebene durchgeführt werden.

6.2.1 Ausschlussbeweis 1 für Softwareänderungen

Gemäß dem ersten Ausschlussbeweis (AB I: Ergebnisneutralität) kann ein Test t_i von der Menge der durchgeführten Tests T ausgeschlossen werden, wenn ein formaler Beweis existiert, der demonstriert, dass die Änderung A keinen Einfluss auf das Ergebnis R_i des Tests t_i hat. Die Einflussfreiheit ist dabei dann gegeben, wenn das Testergebnis $R(t_i, A)$ innerhalb des Vertrauensbereichs des Testergebnisses $R(t_i)$ liegt. Auf Softwareänderungen kann dies wie folgt übertragen werden:

AB I (Softwareänderung):

Wenn nachgewiesen werden kann, dass eine Funktion F innerhalb ihrer variierbaren Parameter im Testfall t_i nie aktiv werden kann, hat die Änderung dieser Funktion keine Auswirkung auf den Testfall. Der Testfall t_i kann daher exkludiert werden.

Die Durchführung dieses Ausschlussbeweises erfolgt durch einen Vergleich der Einsatzgrenzen des Tests t_i mit dem Arbeitsbereich der Funktion F . Die Einsatzgrenzen stellen dabei die Randbedingungen dar, unter denen der Test durchgeführt wird. Da an dieser Stelle Tests im Gesamtfahrzeug betrachtet werden, sind dies alle durch die Testbeschreibung festgelegten Parameter, die die Fahrsituation beschreiben. Eine häufig verwendete Unterteilung dieser Parameter erfolgt in Fahrer, Fahrzeug und Umwelt.¹⁴⁷ Beispiele sind:

¹⁴⁷ Benda, von et al.: Häufigkeit von Verkehrssituationen, 1983; Fastenmeier: Autofahrer Verkehrssituationen, 1995, S. 48 ff; Reichart: Menschliche Zuverlässigkeit, 2001, S. 52; Domsch et al.: Referenzfahrsituationen, 2008, S. 7 nach Weitzel: Objektive Bewertung der Kontrollierbarkeit, 2013, S. 34

- Fahrer¹⁴⁸: Versuchingenieure, Testfahrer
- Fahrzeug: Beladung, Geschwindigkeit, Beschleunigung
- Umwelt: Reibwert, Außentemperatur.

In Anlehnung an die Bestimmung der Testeinsatzgrenzen, markiert der Arbeitsbereich einer Funktion F die definierten Voraussetzungen, unter denen diese Funktion aktiv ist. Hier werden somit die Situationsparameter festgelegt, unter denen die Funktion arbeitet.

Beispiel

In Abbildung 6-2 sind beispielhafte Arbeitsbereiche von fünf Funktionen (A, B, C, D, E) dargestellt. Zur Definition der Arbeitsbereiche dienen hier die Parameter Reibwert μ und Geschwindigkeit v . Eine geänderte Funktion F ist nur bei Geschwindigkeiten von $v > 30$ km/h und auf Hochreibwert aktiv. Wenn nachgewiesen werden kann, dass diese Voraussetzungen immer eingehalten werden, können alle Tests, deren Einsatzgrenzen außerhalb dieses Arbeitsbereichs liegen, bei der Überprüfung der Änderung entfallen.

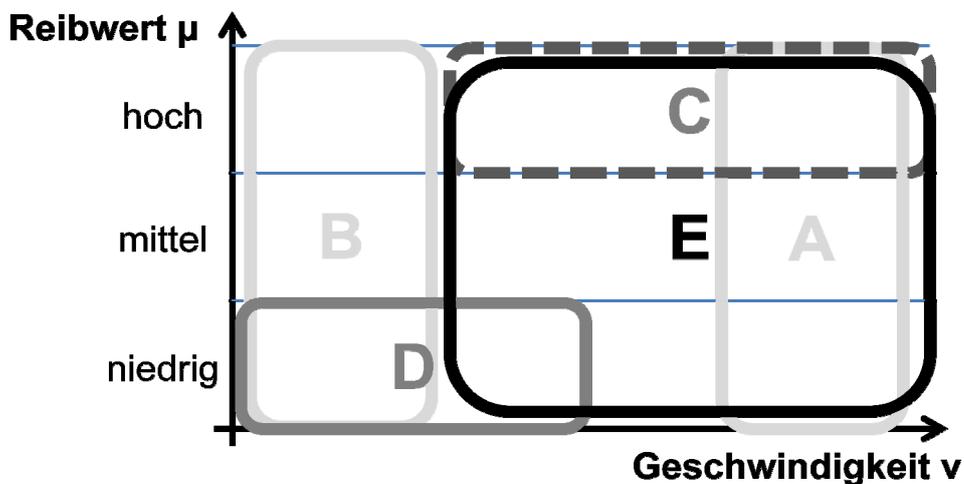


Abbildung 6-2: Definition von Arbeitsbereichen

Grundlage für den Ausschlussbeweis bildet ein Nachweis der Einhaltung des Arbeitsbereichs der Funktion F . Denn nur wenn verifiziert werden kann, dass diese Voraussetzungen unter allen Randbedingungen erfüllt sind, dürfen Testfälle exkludiert werden, deren Einsatzraum außerhalb des Arbeitsbereichs der geänderten Funktion F liegt. Der

¹⁴⁸ Der Begriff Fahrer wird in dieser Arbeit geschlechtsneutral verwendet und umfasst sowohl Fahrerinnen als auch Fahrer.

Beweis kann mit Hilfe von Methoden der Softwaretechnik erbracht werden. Dazu sind grundsätzlich zwei Lösungen denkbar. Bei dynamischen Tests wird die zu testende Funktion mit konkreten Eingabewerten versehen und ausgeführt.¹⁴⁹ Aufgrund des Stichprobencharakters von dynamischen Tests, sind vollständige Aussagen schwer generierbar.¹⁴⁹ Für den erforderlichen Nachweis muss die Funktion F unter allen im Fahrbetrieb auftretenden Situationen überprüft werden, z.B. in einer SiL- oder HiL-Umgebung. Wenn dabei die definierten Voraussetzungen für die Aktivierung immer eingehalten werden, ist der Nachweis erbracht. Der Aufwand für diesen sog. erschöpfenden Test¹⁵⁰ ist bei der Größe des Gesamtkollektivs der möglichen Fahrsituationen im Kundenbetrieb als sehr hoch einzuschätzen, so dass diese grundsätzliche Lösungsmöglichkeit keine praktische Bedeutung erlangen wird. Die zweite Lösung bilden statische Verifikationsverfahren. So versucht der formale Korrektheitsbeweis einen Nachweis der Konsistenz zwischen einer Software und ihrer Spezifikation mit formalen Mitteln zu erbringen.¹⁵¹ Ein anderer Ansatz ist der symbolische Test, bei dem der Quellcode der zu untersuchenden Funktion mit allgemeinen symbolischen Werten in einer künstlichen Umgebung ausgeführt wird.¹⁵¹ Diese Verfahren gestatten es – falls bestimmte Voraussetzungen erfüllt sind - vollständige Aussagen zu erhalten.¹⁵² Somit wird ein Nachweis des Arbeitsbereiches ermöglicht.

Um das Konzept effizient zu nutzen, können Gruppen von Testfällen zusammengefasst werden (z.B. Hochreibwert, Niedrigreibwert oder Geschwindigkeitsbereiche). Dann muss lediglich ein Vergleich mit der Gruppe erfolgen und nicht jeder Testfall einzeln exkludiert werden.

6.2.2 Ausschlussbeweis 1 für Hardwareänderungen

Der erste Ausschlussbeweis (AB I: Ergebnisneutralität) kann auf Hardwareänderungen wie folgt übertragen werden:

AB I (Hardwareänderung):

Wenn nachgewiesen werden kann, dass ein physikalischer Parameter d im Testfall t_i keine Auswirkung hat, hat die Änderung dieses Parameters keine Auswirkung auf den Testfall t_i . Er kann somit exkludiert werden.

¹⁴⁹ Liggesmeyer: Software-Qualität, 2009, S. 39

¹⁵⁰ Liggesmeyer: Software-Qualität, 2009, S. 37

¹⁵¹ Liggesmeyer: Software-Qualität, 2009, S. 335 ff.

¹⁵² Liggesmeyer: Software-Qualität, 2009, S. 44

Auf Änderungen der Fahrzeughardware ist der Ansatz des Vergleichs der Testeinsatzgrenzen mit dem Arbeitsbereich nur teilweise übertragbar, da vielen Fahrzeugparametern keine definierten Arbeitsbereiche zuzuordnen sind (z.B. Karosserie, Reifen, Rad-aufhängung). Für die Umsetzung des ersten Ausschlussbeweises (AB I: Ergebnisneutralität) wird daher ein anderes Konzept vorgeschlagen, das im Folgenden vorgestellt wird. Die zugrundeliegende Systemvorstellung geht dabei auf Feldhusen und Grote¹⁵³ zurück (siehe Abbildung 6-3). Demnach kann jedes technische Artefakt als System aufgefasst werden, das durch Eingangs- und Ausgangsgrößen mit seiner Umgebung in Verbindung steht. Grundsätzlich werden drei Eingangs- und Ausgangsgrößen unterschieden. Dies sind Energie, Stoff und Signal.

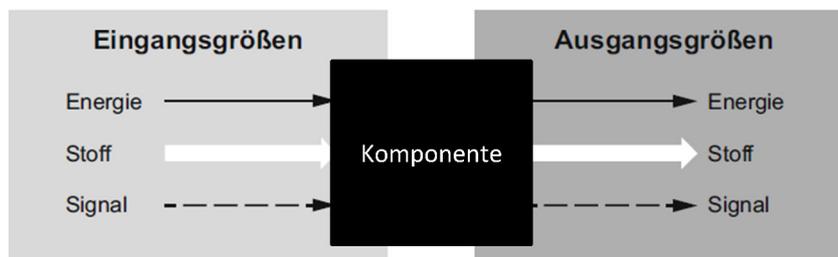


Abbildung 6-3: Systemvorstellung nach Feldhusen und Grote

Eine Änderung eines physikalischen Parameters eines technischen Systems hat demnach eine Auswirkung auf den Umsatz von Energie, Stoff und/oder Signal. Zur Durchführung des Ausschlussbeweises werden nun die Auswirkungen der Änderung des physikalischen Parameters d auf den Energie-, Stoff- und Signalumsatz überprüft. Dazu wird eine endliche Anzahl an Auswirkungskategorien je Klasse (Energie, Stoff, Signal) definiert. Mithilfe von Systembetrachtungen, die die Eingangs- und Ausgangsgrößen einer Komponente analysieren, können alle Auswirkungskategorien ausgeschlossen werden, die von der Änderung nicht betroffen sind. Ein Test t_i kann dann von der Menge der durchzuführenden Tests T ausgeschlossen werden, wenn für alle untersuchten Auswirkungskategorien dieses Tests die Auswirkungsfreiheit begründet werden kann.

Beispiel

Im Rahmen des Absicherungsprozesses wird ein Fahrwerkklager geändert. Über eine Energiebetrachtung an den Systemgrenzen kann nachgewiesen werden, dass am System Fahrwerkklager kein elektrischer Energieumsatz stattfindet. Alle Tests, die lediglich elektrische Auswirkungen überprüfen, können daher exkludiert werden.

¹⁵³ Feldhusen, Grote: Pahl/Beitz Konstruktionslehre, 2013, S. 240

Die Grundlage für den Ausschlussbeweis bildet eine Definition einer endlichen Anzahl an Auswirkungskategorien für die Umsatzklassen Energie, Stoff und Signal. Eine solche Klassifizierung sollte einerseits detailliert genug sein, um die unterschiedlichen Auswirkungskategorien, die in Fahrzeugtests überprüft werden, abzudecken, andererseits aber nicht zu umfangreich ausfallen, um den Ausschlussbeweis noch effizient durchführen zu können. In der Literatur sind unterschiedliche Konzepte zu finden, die Auswirkungskategorien präzisieren (z.B. Szykman et al.¹⁵⁴, Stone und Wood¹⁵⁵, Altshuller¹⁵⁶, Hirtz et al.¹⁵⁷). Hirtz et al.¹⁵⁸ zeigen einen Ansatz, der die Ergebnisse früherer Konzepte zusammenführt und in die drei Umsatzklassen Energie, Stoff und Signal einteilt. Das resultierende Klassifikationsschema („functional basis“) kann als Grundlage für den Ausschlussbeweis herangezogen werden. Hirtz et al. unterscheiden folgende Kategorien:

- Energie: menschlich, akustisch, biologisch, chemisch, elektrisch, elektromagnetisch, hydraulisch, magnetisch, mechanisch, pneumatisch, radioaktiv/ nuklear, thermisch
- Stoff: Mensch, Gas, Flüssigkeit, Festkörper, Plasma, Gemisch¹⁵⁹
- Signal: Status, Regelsignal¹⁶⁰

Das System wurde bei Ford zur Modellierung von komplexen Funktionen angewandt und hat sich in der Praxis bewährt¹⁶¹, weshalb in dieser Arbeit darauf zurückgegriffen wird. Für den Ausschluss von Testfällen können dann z.B. für jede Wechselwirkungskategorie Checklisten angelegt werden, auf denen Fragen vermerkt sind, die ein Tester bzw. Entwickler verneinen muss, um diese Kategorie ausschließen zu können.

Das zuvor dargestellte Beispiel zum Fahrwerkklager zeigt, dass das Konzept grundsätzlich funktioniert. Es führt aber auch die Grenzen der Methode vor Augen. Der Ausschluss von weiteren Testfällen ist in diesem Beispiel sehr schwierig. Um das Konzept effizienter zu nutzen, können für einzelne Komponenten bestimmte Wechselwirkungen

¹⁵⁴ Szykman et al.: The representation of function in computer-based design, 1999

¹⁵⁵ Stone, Wood: Development of a functional basis for design, 1999

¹⁵⁶ Altshuller: Creativity as an exact science, 1984

¹⁵⁷ Hirtz et al.: A functional basis for engineering design, 2002

¹⁵⁸ Hirtz et al.: A functional basis for engineering design, 2002

¹⁵⁹ Stoffflüsse sind im Fahrzeugbereich prinzipiell denkbar (z.B. Leckage im Hydrauliksystem), spielen aber eine untergeordnete Rolle und können daher auch vernachlässigt werden.

¹⁶⁰ Da bei diesem Beweis lediglich Hardwareänderungen betrachtet werden, können Einflüsse auf Signale vernachlässigt werden.

¹⁶¹ Hirtz et al.: A functional basis for engineering design, 2002

immer bejaht werden. In dem gezeigten Beispiel sind dies die mechanischen Wechselwirkungen. Dadurch werden die Testfälle, die mechanische Auswirkungen haben, immer automatisch inkludiert und der Aufwand für die erneute Auswahl dieser Testfälle bei Änderungen entfällt. Damit wird der Beweis effizienter. Außerdem können den einzelnen Komponenten Komplexitätsmarker zugeordnet werden. Je komplexer eine Komponente ist, desto schwieriger gestaltet sich die Exklusion von Testfällen. Mit Hilfe der Komplexitätsmarker können die Testingenieure dann gezielte Schwerpunkte bei der Suche nach Testfällen für die Exklusion setzen.

6.2.3 Ausschlussbeweis 2 für Software- und Hardwareänderungen

Gemäß dem zweiten Ausschlussbeweis (AB II: Ergebnisgleichheit) kann ein Test von der Menge der durchgeführten Tests ausgeschlossen werden, wenn eine Kombination anderer Tests existiert, deren Ergebnisse zusammengenommen den Einfluss der Änderung mindestens genauso gut zeigen wie das Ergebnis des Tests.

AB II (Software- und Hardwareänderung):

Wenn nachgewiesen werden kann, dass die Ergebnisse $R_j \dots R_n$ der in der Testmenge verbleibenden Tests $t_j \dots t_n$ ($t_j \dots t_n \in T$) mögliche (qualitätsverschlechternde) Auswirkungen der Änderung A mindestens genauso deutlich zeigen, wie das Ergebnis R_i des Tests t_i , kann der Testfall t_i ausgeschlossen werden.

Die Durchführung dieses Ausschlussbeweises erfolgt durch einen Vergleich der Testziele des Tests t_i mit den Testzielen der in der Testmenge verbleibenden Tests ($T - \{t_i\}$). Der Test t_i kann ausgeschlossen werden, wenn jedes Testziel von t_i durch einen anderen Testfall, der durchgeführt wird, abgedeckt wird. Die Testfälle, die als Referenz für den Ausschluss von t_i herangezogen werden, dürfen dann allerdings nicht weiter exkludiert werden, sie erfahren damit eine Inklusion.

Beispiel

Eine exemplarische Testmenge $T = \{t_1, t_2, t_3, t_4\}$ mit den Testzielen $z_1 - z_5$ ist in Abbildung 6-4 dargestellt. Der Test t_3 besitzt zwei Testziele (z_2 und z_4). Beide Testziele werden auch durch Test t_2 abgedeckt. Der Test t_3 kann damit von der Testmenge T ausgeschlossen werden. Der Referenztest t_2 darf dann nicht weiter exkludiert werden.

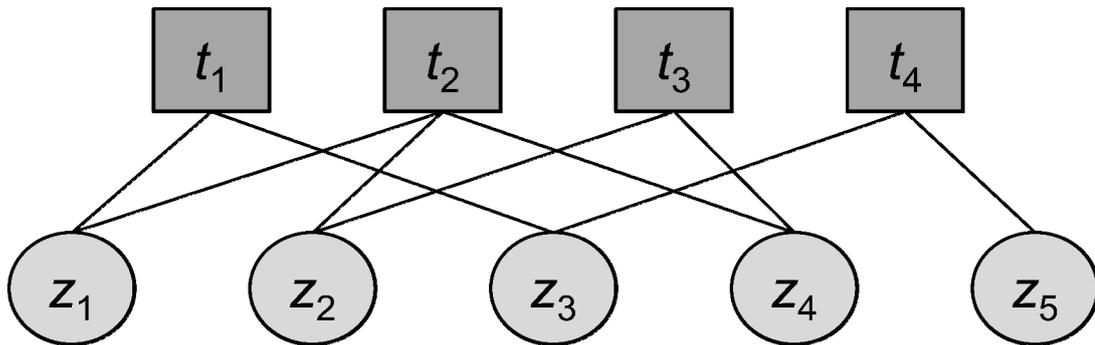


Abbildung 6-4: Beispielhafte Testmenge für AB II

Die Durchführung des zweiten Ausschlussbeweises kann effizient mithilfe von Methoden aus der Softwaretechnik umgesetzt werden. Hierzu bieten sich Test Suite Reduction Methods an. Diese versuchen aus einer Testmenge eine minimale Untermenge zu bilden, die bestimmte Anforderungen erfüllt. Das Optimierungsproblem ist in der Softwaretechnik als Mengenüberdeckungsproblem („minimum set cover problem“) bekannt. Ein Beispiel für einen Lösungsansatz ist der H-Algorithmus von Harrold et al.¹⁶², der das Ziel verfolgt, bei der Testauswahl die maximale Codeabdeckung bei minimaler Testanzahl zu erhalten, d.h. jeder Teil des Codes soll mindestens einmal ausgeführt werden. Andere Konzepte sind z.B. bei Parsa et al.¹⁶³ oder Kichigin¹⁶⁴ zu finden. Übertragen auf den o.g. Ausschlussbeweis kann damit die minimale Untermenge von T bestimmt werden, die alle Testziele mindestens einmal abdeckt. Im Beispiel in Abbildung 6-4 decken die Testfälle t_2 und t_4 alle Testziele ab. Sie bilden die minimale Untermenge, die durchgeführt werden muss, um jedes Testziel einmal abzurufen. Zusätzlich zur Minimierung der Gesamtanzahl an Tests können noch weitere Optimierungskriterien berücksichtigt werden (z.B. Kosten, Ausführungszeit oder Sicherheitsrelevanz der Test) (siehe z.B. Baller et al.¹⁶⁵). Darüber hinaus bieten die genannten Methoden noch weiteres Potenzial für Aufwandsreduktionen in der Gesamtfahrzeugabsicherung. So ist damit z.B. auch eine Optimierung des Gesamttestkatalogs möglich. Dadurch könnten Redundanzen, d.h. mehrfach überprüfte Testziele, identifiziert und damit die Testkataloge insgesamt verkleinert werden.

¹⁶² Harrold et al.: A Methodology for Controlling the Size of a Test Suite, 1993

¹⁶³ Parsa et al.: A New Method to Test Suite Reduction Based on Cluster Analysis, 2009

¹⁶⁴ Kichigin: A Method for Test Suite Reduction for Regression Testing, 2010

¹⁶⁵ Baller et al.: Multi-Objective Test Suite Optimization for Incremental Product Family Testing, 2014

6.3 Fazit

In den vorangegangenen Abschnitten wurde deduktiv ein Ansatz zur exklusionsbasierten Testauswahl abgeleitet und beispielhaft konkretisiert. Die Ergebniserwartung der beschriebenen exklusionsbasierten Testauswahlmethode in Form eines ROC-Diagramms (vgl. Abschnitt 4.4.1) ist in Abbildung 6-5 dargestellt.

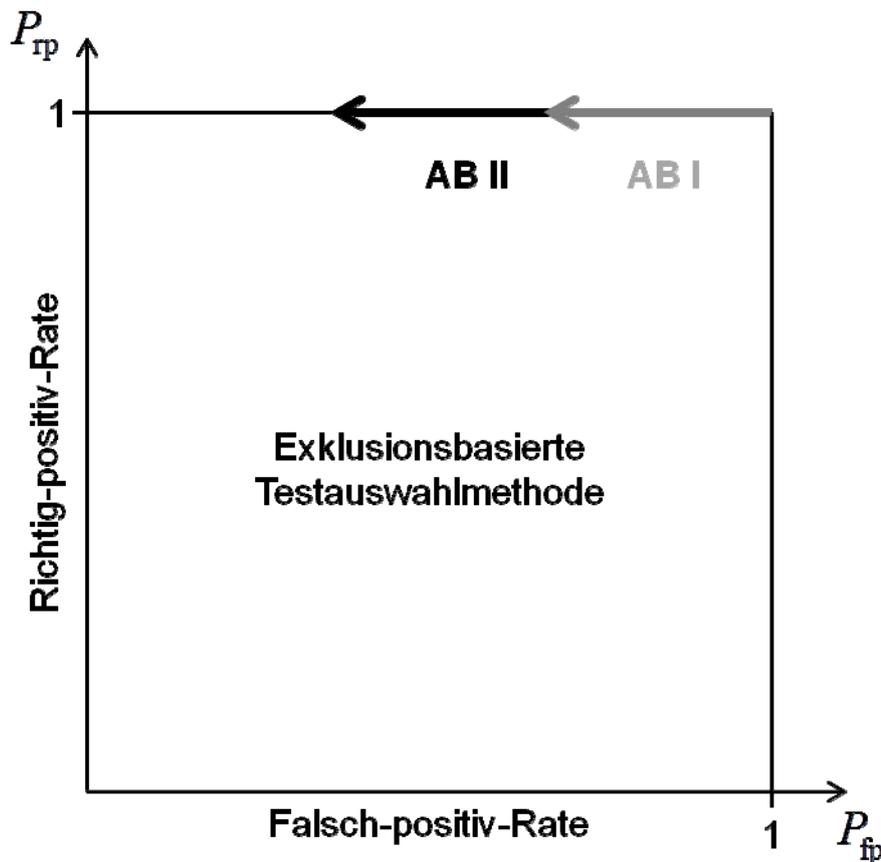


Abbildung 6-5: Ergebniserwartung für die exklusionsbasierte Testauswahlmethode

Die beiden Ausschlussbeweise eliminieren Tests von der Menge der durchzuführenden Tests. Der große Vorteil der gezeigten exklusionsbasierten Testauswahl ist die hohe Vollständigkeit. Die resultierende Testmenge $T_{\text{red},2}$ ist immer sicher (i.S.v. keine falsch negativen Tests), da nur obsolete Tests ausgeschlossen werden und so alle notwendigen Tests in der Menge der durchgeführten Tests enthalten bleiben. Falsch negative Ergebnisse werden dadurch vermieden. Dies macht eine Anwendung im rechtlich relevanten Bereich der Gesamtfahrzeugabsicherung möglich.

Durch die Exklusion von Falsch-Positiven wird die Fehlerrate reduziert. Eine Fehlerrate von null kann allerdings nur erreicht werden, wenn alle Tests, die nicht notwendig sind, exkludiert werden. Wie die exemplarische Umsetzung zeigt, ist dieser Ausschlussbeweis in der Praxis jedoch sehr aufwändig, da ein Testfall nur ausgeschlossen werden kann, wenn er nachweislich in keiner Wechselwirkung mit der Änderung steht. Daraus folgt, dass die Methode mit einem sehr hohen Selektionsaufwand verbunden ist. Jede Unsicherheit in der Beurteilung führt darüber hinaus automatisch zu einer Inklusion des Tests. Dies hat zur Folge, dass in der Menge der durchgeführten Tests auch Testfälle enthalten sein können, die eigentlich nicht notwendig sind, für die jedoch der Beweis der Wechselwirkungsfreiheit nicht möglich war (falsch positive Tests). Der ideale Punkt ($P_{pf} = 0$, $P_{rp} = 1$) kann also nur schwer erreicht werden.

7 Entwicklung eines Gesamtkonzeptes für die Testauswahl

Nachdem in den vorangegangenen Kapiteln die spezifischen Vor- und Nachteile der beiden prinzipiellen Lösungskonzepte anhand exemplarischer Umsetzungen diskutiert wurden, erfolgt in diesem Kapitel die Entwicklung eines Gesamtkonzeptes für die Testauswahl bei Änderungen in der Gesamtfahrzeugabsicherung.

7.1 Vergleich von Inklusion und Exklusion

Tabelle 7-1 stellt die Eigenschaften der beiden Methoden zusammenfassend gegenüber.

Tabelle 7-1: Vergleich von Inklusions- und Exklusionsanalyse

	Inklusionsanalyse	Exklusionsanalyse
Prinzip	Inklusion von Testfällen, die von der Änderung betroffen sind.	Exklusion von Testfällen, die von der Änderung nicht betroffen sind
Vollständigkeit	Eine hohe Vollständigkeit kann nur erreicht werden, wenn alle Falsch-Negativen (f_n) identifiziert und in Richtig-Positive (r_p) überführt werden. Dazu müssen alle Wechselwirkungen der Änderung ermittelt und Testfällen zugeordnet werden. Entscheidend für die Qualität der Testauswahl ist die Güte des zugrundeliegenden Systemmodells.	Die Testauswahl ist sicher i.S.v. Vollständigkeit. Falsch negative Ergebnisse werden vermieden (d.h. $f_n = 0$), da nur solche Testfälle ausgeschlossen werden, für die Belege vorliegen, dass sie in keiner Wechselwirkung zur Änderung stehen.
Fehlerrate	Falsch-Positive können nicht ausgeschlossen werden, da es unmöglich ist, effizient alle fehleraufdeckenden Tests zu	Falsch-Positive sind sehr wahrscheinlich. Alle Testfälle, für die kein Beweis der Wechselwirkungsfreiheit

	finden und daher immer eine Übermenge an Testfällen (sog. modifikationsdurchlaufende Tests) ermittelt wird.	erbracht werden kann, müssen durchgeführt werden.
Selektionskosten	Die Kosten der inklusionsbasierten Testauswahl beinhalten Kostenanteile, die durch Falsch-Negative (f_n) und durch Falsch-Positive (f_p) entstehen, da beide Kategorien nicht prinzipiell vermieden werden.	Die Kosten der exklusionsbasierten Auswahl bestehen nur aus den Kostenanteilen, die durch Falsch-Positive (f_p) verursacht werden, weil Falsch-Negative (f_n) ausgeschlossen sind.
Selektionsaufwand	Der Selektionsaufwand kann sehr gering ausfallen, falls die zugrundeliegende Informationsquelle verfügbar ist, d.h. nicht speziell für diesen Zweck ermittelt werden muss, und der Prozess der Testauswahl automatisiert erfolgt (siehe Beispiele für effiziente Methoden aus dem Stand der Technik).	Der Selektionsaufwand ist nach derzeitigem Stand hoch.

Die Konkretisierungen der beiden prinzipiellen Lösungsmöglichkeiten für die Testauswahl zeigen, dass die Annahme einer idealen Selektionsqualität (SQ) nicht haltbar ist. Beiden Ansätzen ist gemein, dass sie keine Fehlerrate von null besitzen. Bei der Inklusionsanalyse entstehen falsch positive Ergebnisse (f_p) dadurch, dass es keine Möglichkeit gibt, fehleraufdeckende Tests effizient zu ermitteln. Es wird daher immer eine Übermenge von Tests bestimmt, die alle fehleraufdeckenden Tests enthält (sog. modifikationsdurchlaufende Tests). Bei der Exklusionsanalyse sind falsch positive Tests ebenfalls sehr wahrscheinlich, da alle Testfälle, für die kein Beweis der Wechselwirkungsfreiheit erbracht wird, durchgeführt werden – auch wenn diese nicht notwendig sind. Die Exklusionsanalyse besitzt gegenüber der Inklusionsanalyse den großen Vorteil, dass die Testauswahl sicher ist. Falsch negative Ergebnisse werden vermieden (d.h. $f_n = 0$), da nur solche Testfälle ausgeschlossen werden, für die Belege vorliegen, dass sie in keiner Wechselwirkung zur Änderung stehen. Bei der Inklusionsanalyse ist die Vollständigkeit hingegen stark von der Güte des zugrundeliegenden Modells zur Ermittlung der Wechselwirkungen der Änderung abhängig. Um Änderungen auf Gesamtfahrzeugebene nachzuverfolgen, wird ein Modell benötigt, das alle Wechselwirkungen vollständig und korrekt abbildet. Aufgrund der hohen Gesamtsystemkomplexität ist es jedoch

sowohl schwierig, ein solches Modell in ausreichender Genauigkeit zu entwickeln, als auch dieses zu validieren. Das im Sinne einer hohen Selektionsqualität sowie geringen Selektionsergebniskosten zu bevorzugende Konzept ist demnach die Exklusionsanalyse. Großer Nachteil dieses Konzeptes ist jedoch die geringe Effizienz der Testauswahl, da pro Testfall immer alle Wechselwirkungen mit der Änderung belegbar ausgeschlossen werden müssen. Im Hinblick auf einen geringen Selektionsaufwand ist die Inklusionsanalyse zu bevorzugen, da im Stand der Technik bereits viele Methoden existieren, die dieses Konzept effizient umsetzen. Das Gesamtkonzept wird daher so ausgelegt, dass es den Vorteil der hohen Vollständigkeit der Exklusion mit dem Vorteil der großen Effizienz der Inklusion verbindet. Wie dies umgesetzt werden kann, wird in den folgenden Abschnitten beschrieben.

7.2 Gesamtkonzept für die Testauswahl

Bevor das Gesamtkonzept vorgestellt wird, erfolgt eine Definition zugrundeliegender Begrifflichkeiten (siehe Tabelle 7-2).

Tabelle 7-2: Definitionen für das Gesamtkonzept der Testauswahl

Symbol	Name	Beschreibung
T_W	Weißer Testfälle	Testfälle, bei denen eine Wechselwirkung mit der Änderung sicher ist
T_{HG}	Hellgraue Testfälle	Testfälle, bei denen es Indikatoren gibt, die Hinweise auf Wechselwirkungen mit der Änderung geben
T_{DG}	Dunkelgraue Testfälle	Testfälle, bei denen es keinen Hinweis auf Wechselwirkungen mit der Änderung gibt
T_S	Schwarze Testfälle	Testfälle, bei denen eine Wechselwirkung mit der Änderung ausgeschlossen ist

Als weiße Testfälle, T_W , werden Tests bezeichnet, bei denen eine Wechselwirkung mit der Änderung sicher ist. Das sind die sog. fehleraufdeckenden Tests, die bei der Inklusionsanalyse ermittelt werden sollen.¹⁶⁶ Wie die vorangegangenen Ausführungen in Kap. 5 zeigen, handelt es sich hierbei um eine ideale Testfallkategorie, für die nach Rothermel jedoch keine effiziente Methode existiert, mit der diese Testfälle identifiziert wer-

¹⁶⁶ Vgl. Abschnitt 5.1

den können. Bei hellgrauen Testfällen, T_{HG} , handelt es sich um solche Tests, bei denen es Indikatoren gibt, die Hinweise auf Wechselwirkungen mit der Änderung geben. Bei diesen Indikatoren kann es sich z.B. um Abhängigkeiten im Quelltext oder um Wechselbeziehungen im Fahrzeugmodell handeln. Das Ergebnis heutiger Verfahren der Inklusionsanalyse beinhaltet genau diese Testfälle. Da bei den hellgrauen Testfällen eine hohe Wahrscheinlichkeit für eine Abhängigkeit zur Änderung besteht, lassen sich die Testfälle – wenn überhaupt – nur mit hohem Aufwand über eine Exklusionsanalyse ausschließen. Dunkelgraue Testfälle, T_{DG} , hingegen zeigen keine Hinweise auf Wechselwirkungen. Dies betrifft die Testfälle, die bei der Inklusionsanalyse nicht als notwendig identifiziert werden. Bei diesen Testfällen liegt ein hohes Potenzial für einen Ausschluss durch Exklusion bei vertretbarem Aufwand vor. Die letzte Kategorie umfasst die schwarzen Testfälle, T_S . Bei diesen Tests handelt es sich um Beispiele, bei denen eine Wechselwirkung sicher ausgeschlossen ist. Dies sind die Tests, die mit Hilfe einer Exklusionsanalyse als nicht notwendig identifiziert wurden, da alle Wechselwirkungen mit der Änderung belegbar ausgeschlossen werden.

Das Gesamtkonzept für die Testauswahl ergibt sich aus einer Effizienzbetrachtung. Da die Inklusionsanalyse zwar nicht sicher i.S.v. Vollständigkeit, aber sehr effizient umsetzbar ist, wird sie zur Vorselektion von Testfällen für die Exklusion eingesetzt, deren Selektion sicher, aber aufwändig ist. So werden die spezifischen Vorteile beider Ansätze kombiniert. Es entsteht daher das in Abbildung 7-1 dargestellte Gesamtkonzept.

Ausgangsbasis für die Testauswahl ist die Testmenge T , d.h. die Menge aller Tests, die zur Absicherung des Gesamtsystems herangezogen werden. Im ersten Auswahlschritt wird die Inklusionsanalyse durchgeführt. Dazu eignen sich besonders die Methoden aus dem Stand der Technik, die sich in heutige Prozesse einfach integrieren lassen. Dies sind z.B. die vielfach beschriebenen RTS-Techniken, die auf einer Analyse des Quelltextes von Systemsoftware aufbauen.¹⁶⁷ Die Inklusion identifiziert alle Tests, bei denen eine Wechselwirkung mit der Änderung wahrscheinlich ist. Sie dient damit der effizienten Vorselektion der Testfälle für die Exklusion. Für die durch die Inklusion identifizierten Testfälle gibt es Hinweise auf Wechselwirkungen, so dass diese Testfälle – wenn überhaupt – nur mit hohem Aufwand auszuschließen sind (hellgraue Testfälle T_{HG}). Für diese Tests lohnt sich die Exklusion vermutlich nicht. Sie werden daher direkt in die Menge der durchzuführenden Tests T' transferiert. Die Testfälle, bei denen die Inklusionsanalyse keine Hinweise auf Wechselwirkungen ermittelt hat, bieten somit ein hohes Potenzial für den Ausschluss (dunkelgraue Testfälle T_{DG}). Sie bilden die Ausgangsbasis für die in Kap. 6.2 beschriebene Exklusionsanalyse, die den zweiten Auswahlschritt des Testkonzeptes darstellt. Hier wird eine Exklusion von Testfällen durchgeführt, bei denen Wechselwirkungen der Änderung nachweislich nicht vorhanden sind,

¹⁶⁷ Vgl. Kap. 3.4.2 und Kap. 5.2

sog. schwarze Testfälle T_S . Tests, für die dieser Ausschluss nicht möglich ist, werden durchgeführt.

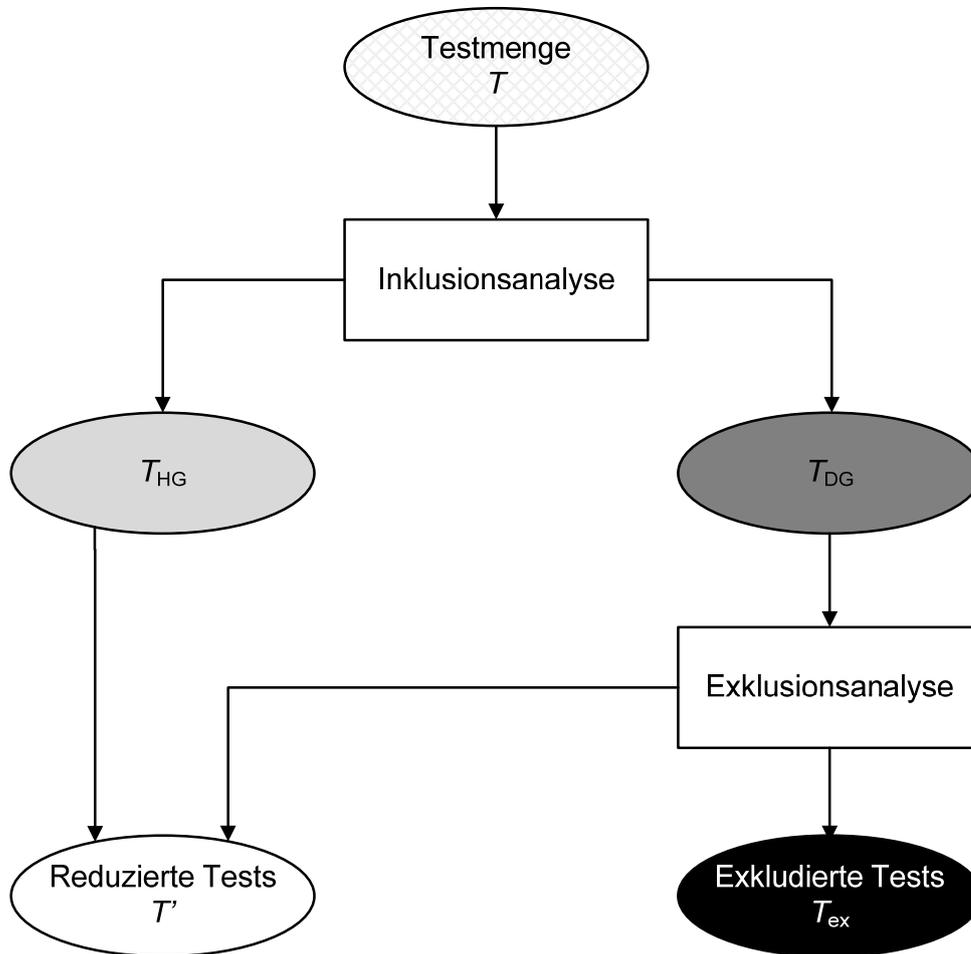


Abbildung 7-1: Gesamtkonzept der Testauswahl

Das Testkonzept ermöglicht damit eine zweistufige Auswahl von Testfällen. Eine Einschränkung des Konzeptes ergibt sich in den Fällen, in denen die Änderung einen Testfall erfordert, der nicht im Ausgangstestkatalog enthalten ist. Solche Fälle können mit der Methodik nicht identifiziert werden. Dies ist jedoch nicht relevant solange nur Änderungen betrachtet werden, die keine Funktionserweiterungen darstellen.

7.3 Ergebniserwartung

Ziel der Testauswahl bei Änderungen in der Gesamtfahrzeugabsicherung ist einerseits, die Testfälle zu inkludieren, die notwendig zur Überprüfung der Änderung sind. Dies sind die „wahren“ Richtig-Positiven (RP_w). Andererseits sind die Tests zu exkludieren,

die zur Kontrolle der Änderung nicht notwendig sind. Dies sind die „wahren“ Richtig-Negativen (RN_w). Im Prinzip stellt die Testauswahl daher eine Klassifizierung der Tests in die beiden genannten Gruppen dar. Sie kann graphisch anhand eines Verteilungsdiagramms dargestellt werden, bei dem die Anzahl der „wahren“ Richtig-Positiven (RP_w) auf der Ordinate und die Anzahl der „wahren“ Richtig-Negativen (RN_w) auf der Abszisse aufgetragen wird (siehe Abbildung 7-2).

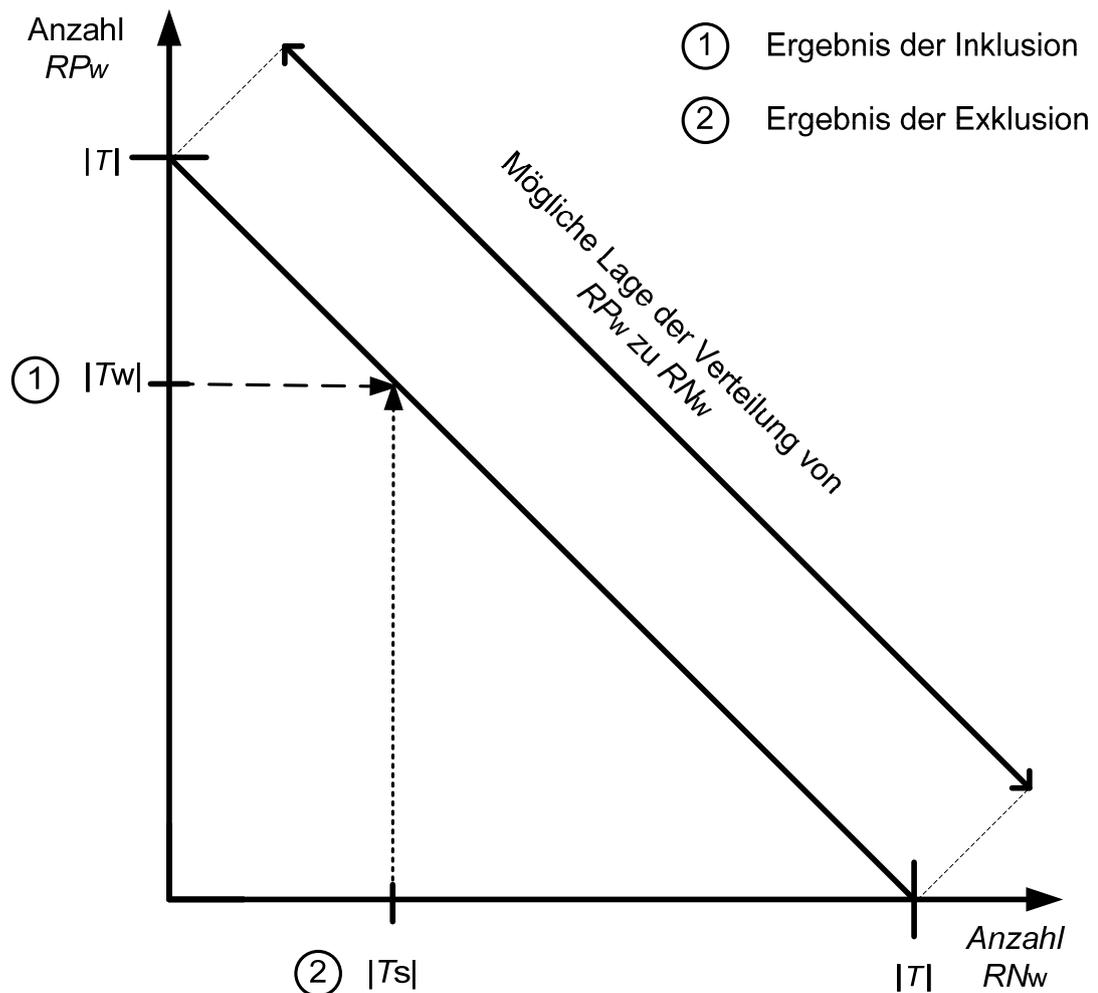


Abbildung 7-2: Ergebniserwartung bei idealer Inklusion und Exklusion

Die maximale Anzahl von „wahren“ Richtig-Positiven beträgt $|T|$. In diesem Fall (Punkt $(0, |T|)$ in Abbildung 7-2) sind alle Tests zur Überprüfung der Änderung notwendig. Es ist jedoch auch möglich, dass alle Tests nicht notwendig sind, dann ergibt sich die maximale Anzahl von „wahren“ Richtig-Negativen als $|T|$ (Punkt $(|T|, 0)$ in Abbildung 7-2). Die Verteilung von RP_w zu RN_w ist zu Beginn der Selektion unbekannt. Der Punkt, der die wahre Lage kennzeichnet, (RN_w, RP_w) , liegt allerdings auf der Geraden durch die Punkte $(0, |T|)$ und $(|T|, 0)$. Eine ideale Inklusion ermittelt alle weißen Testfälle (T_w). Die Anzahl der weißen Testfälle entspricht dann der Anzahl der „wahren“ Richtig-

Positiven und es ergibt sich automatisch die Anzahl der „wahren“ Richtig-Negativen aus der Differenz ($|T|-|T_W$). Bei idealer Exklusion werden alle schwarzen Testfälle (T_S) identifiziert. Sie entsprechen der Anzahl der „wahren“ Richtig-Negativen. Auch hieraus ergibt sich dann automatisch die Anzahl der „wahren“ Richtig-Positiven. Ideal betrachtet sind daher Inklusion und Exklusion gleichwertig.

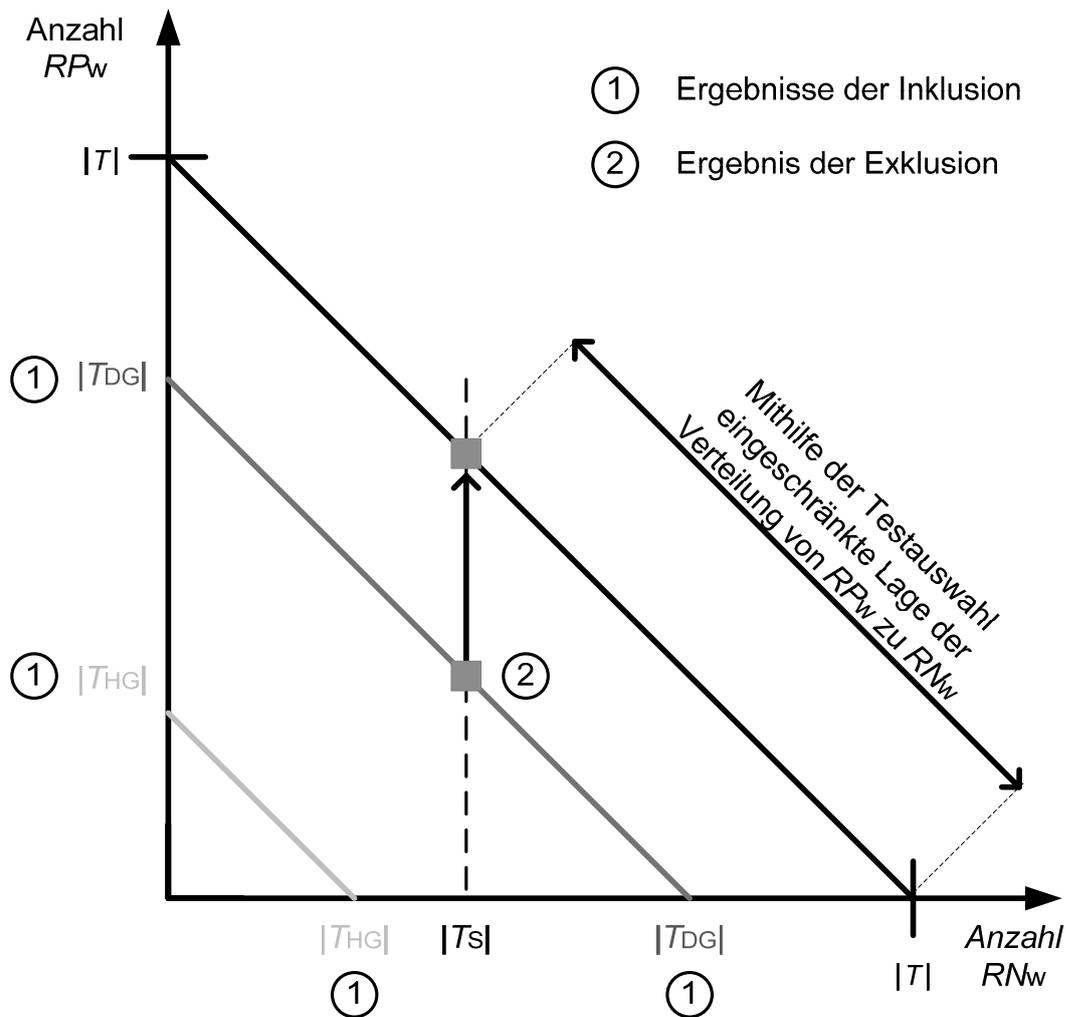


Abbildung 7-3: Ergebniserwartung beim vorgeschlagenen Konzept für die Testauswahl

Die Inklusionsanalyse im vorgeschlagenen Gesamtkonzept unterteilt die Gesamtmenge T in zwei Gruppen (hellgraue und dunkelgraue Linie in Abbildung 7-3). Die erste Gruppe bilden die hellgrauen Testfälle (T_{HG}), bei denen es Indikatoren gibt, die Hinweise auf Wechselwirkungen mit der Änderung geben. Diese Tests werden aus Gründen der Effizienz automatisch in die Menge der durchzuführenden Tests übernommen. Sie geben allerdings keinen Hinweis auf die reale Verteilung von RP_w zu RN_w (hellgraue Linie in Abbildung 7-3). Die andere Gruppe umfasst die dunkelgrauen Testfälle (T_{DG} mit $|T_{DG}| = |T| - |T_{HG}|$), bei denen die Inklusionsanalyse keine Indikatoren für Wechselwir-

kungen gefunden hat. Auch diese Testfälle liefern zunächst keine Anhaltspunkte für die reale Verteilung von RP_w zu RN_w (dunkelgraue Linie in Abbildung 7-3). Sie bieten allerdings hohes Potenzial für die Exklusion.

Die dunkelgrauen Testfälle (T_{DG}) werden daher im zweiten Schritt des Testauswahlkonzeptes der Exklusionsanalyse unterzogen. Diese liefert schwarze Testfälle (T_S), bei denen eine Wechselwirkung mit der Änderung ausgeschlossen wird. Die Anzahl der schwarzen Testfälle ($|T_S|$) legt den Abszissenabschnitt eines Punktes auf der Geraden durch die Punkte $(0, |T_{DG}|)$ und $(|T_{DG}|, 0)$ fest, womit die Verteilung von RP_w zu RN_w eingeschränkt werden kann. Die Lage des Punktes (RN_w, RP_w) lässt sich jedoch auch mit diesem Ansatz nicht exakt bestimmen, da die Exklusionsanalyse nur solche Testfälle identifiziert, für die ein Nachweis der Wechselwirkungsfreiheit erbracht werden kann. Es ist also möglich, dass auch nach der Exklusion noch weitere schwarze Testfälle in der Menge der durchzuführenden Tests T' enthalten sind. Die Begrenzung der Lage der Verteilung von RP_w zu RN_w stellt allerdings einen großen Fortschritt im Vergleich zu Methoden aus dem Stand der Technik dar, die auf dem Inklusionskonzept aufbauen. Sie bestimmen hellgraue und dunkelgraue Testfälle, die jedoch keine Hinweise auf die Lage des Punktes (RN_w, RP_w) liefern.

Das vorgeschlagene Konzept ermöglicht zusammengefasst eine sichere Identifizierung von Testfällen, die bei der Überprüfung von Änderungen in der Gesamtfahrzeugabsicherung entfallen können. Im Vergleich zu Methoden aus dem Stand der Technik kann damit die Verteilung von RP_w zu RN_w zuverlässig eingeschränkt werden. Dem Nachteil des hohen Aufwandes des Exklusionskonzeptes wird mit der Vorselektion der Testfälle durch Inklusionsanalysen aus dem Stand der Technik entgegengewirkt, so dass die Methode insgesamt effizient einsetzbar ist.

8 Validierungsansatz für das Testauswahlkonzept

Im nächsten Schritt erfolgt die Beschreibung eines Validierungsansatzes für das aufgestellte Testauswahlkonzept. Dazu werden zunächst die Untersuchungsziele definiert. Darauf aufbauend erfolgt eine Diskussion der Evaluierungsansätze für jedes Untersuchungsziel.

8.1 Untersuchungsziele

Ziel der Validierung ist, zu untersuchen, ob die Methode dem Einsatzzweck, d.h. der Überprüfung von Änderungen in der Gesamtfahrzeugabsicherung, gerecht wird. Dazu sind drei Validierungsziele zu betrachten. Erstens ist zu prüfen, welche Güte der Testauswahl mit dem vorgeschlagenen Konzept erreicht werden kann. Es ist daher zum einen die Sicherheit der Testauswahl i.S.v. falsch negativ klassifizierten Tests zu ermitteln. Dies ist vor allem vor dem Hintergrund der hohen rechtlichen Relevanz der Gesamtfahrzeugabsicherung (vgl. Abschnitt 2.3 und Abschnitt 3.3) als letzte Teststufe vor der Produktion bzw. Auslieferung an den Kunden von zentraler Bedeutung für die Akzeptanz der Methode. Zum anderen ist es notwendig, den Anteil von falsch positiv eingestuften Tests zu bestimmen, um zu bewerten, welche Fehlerrate bei der Testauswahl erzielt werden kann. Aufgrund des späten Zeitpunktes der Gesamtfahrzeugabsicherung im Entwicklungsprozess (vgl. Abschnitt 2.6) und dem hohen Aufwand, der bei der Absicherung von Änderungen entsteht (vgl. Abschnitt 3.5), ist zweitens zu untersuchen, ob die Methode effizient einsetzbar ist. Dazu ist es erforderlich, die Kosten der Testauswahl detailliert zu erfassen. Ferner kann dann ein Vergleich der Effizienz mit der expertenbasierten, manuellen Testauswahl (vgl. Abschnitt 3.4.1) erfolgen. Da die Gesamtfahrzeugabsicherung sehr heterogen ist (vgl. Abschnitt 2.4) und der Umgang mit Änderungen hier stark von der hohen Systemkomplexität geprägt ist (vgl. Abschnitt 3.5), ist an dritter Stelle zu prüfen, inwieweit die Methode generell für Änderungen in der Gesamtfahrzeugabsicherung nutzbar ist bzw. welche Einschränkungen beim Einsatz bestehen. Dazu sind die Verwendbarkeit der Methode bezüglich verschiedener Arten von Änderungen sowie die Anwendbarkeit auf unterschiedliche Testkategorien der Gesamtfahrzeugabsicherung zu analysieren (vgl. Abschnitt 2.4).

8.2 Evaluierung der Güte der Testauswahl

Auf Basis der vier Kenngrößen Richtig-Positive (r_p), Richtig-Negative (r_n), Falsch-Positive (f_p) und Falsch-Negative (f_n) können folgende Kriterien in Anlehnung an Runkler¹⁶⁸ für die Beurteilung der Güte der Testauswahl herangezogen werden.

Sensitivität (Richtig-positiv-Rate)

Die Sensitivität beschreibt die Richtig-positiv-Rate, d.h. den Anteil der bei der Testauswahl ermittelten notwendigen Tests an der Gesamtheit der notwendigen Tests (vgl. auch Abschnitt 4.4.1). Sie stellt damit die Vollständigkeit der Auswahlmethode dar. Bei einer Sensitivität von $P_{tp} = 1$ wird in Anlehnung an Rothermel¹⁶⁹ von einer sicheren Auswahlmethode gesprochen. Eine Auswahlmethode im Bereich der Absicherung von Änderungen sollte, u.a. aufgrund der rechtlichen Relevanz dieses Prozesses (vgl. Abschnitt 3.3), alle notwendigen Tests erfassen, d.h. es wird eine Sensitivität von 1 gefordert.

$$P_{tp} = \frac{r_p}{r_p + f_n} \quad (8-1)$$

Spezifität (Richtig-negativ-Rate)

Die Spezifität ermittelt den Anteil der nicht durchgeführten Tests an der Gesamtheit der nicht notwendigen Tests, d.h. die Richtig-negativ-Rate. Eine Testauswahlmethode für das Anwendungsgebiet der Änderungsabsicherung auf Gesamtfahrzeugebene sollte möglichst viele der nicht notwendigen Tests ausschließen, um den Gesamtaufwand für den Wiederholungstest gering zu halten. Es wird daher eine möglichst hohe Spezifität gefordert.

$$P_m = \frac{r_n}{r_n + f_p} \quad (8-2)$$

Falsch-negativ-Rate

Die Falsch-negativ-Rate stellt den Anteil der nicht durchgeführten Tests an der Gesamtheit der notwendigen Tests dar. Sie beschreibt damit das Risiko der Auswahlmethode.

¹⁶⁸ Runkler.: Data Mining, 2010, S. 87

¹⁶⁹ Rothermel, Harrold: Analyzing Regression Test Selection Techniques, 1996

Da das Risiko möglichst gering ausfallen soll, wird eine Falsch-negativ-Rate von null gefordert.

$$P_{fn} = \frac{f_n}{r_p + f_n} \quad (8-3)$$

Falsch-positiv-Rate

Die Falsch-positiv-Rate kennzeichnet den Anteil der durchgeführten Tests an der Gesamtheit der nicht notwendigen Tests (vgl. auch Abschnitt 4.4.1). Sie gibt also in negativer Weise an, wie gut nicht notwendige Tests durch die Auswahlmethode vermieden werden. Die Falsch-positiv-Rate sollte möglichst gering sein.

$$P_{fp} = \frac{f_p}{r_n + f_p} \quad (8-4)$$

Positiver Vorhersagewert

Der positive Vorhersagewert entspricht dem Anteil der notwendigen Tests an den ausgewählten Tests. Er beschreibt damit die Präzision der Auswahlmethode. Eine Auswahlmethode sollte möglichst nur Testfälle auswählen, die auch notwendig sind, um den Absicherungsaufwand gering zu halten. Es wird daher ein möglichst hoher positiver Vorhersagewert angestrebt.

$$P_p = \frac{r_p}{r_p + f_p} \quad (8-5)$$

Negativer Vorhersagewert

Der negative Vorhersagewert kennzeichnet den Anteil der korrekt nicht ausgewählten Tests an den nicht ausgewählten Tests. Er beschreibt damit die Trennfähigkeit der Testauswahl. Eine Methode zur Absicherung von Änderung auf Gesamtfahrzeugebene sollte möglichst nur Tests ausschließen, die nicht notwendig sind, d.h. es wird ein negativer Vorhersagewert von 1 gefordert.

$$P_n = \frac{r_n}{r_n + f_n} \quad (8-6)$$

Die oben beschriebenen Gütekriterien setzen voraus, dass der „wahre“ Zustand eines Tests, d.h. „notwendig“ oder „nicht notwendig“, vorliegt oder durch eine andere Me-

thode ermittelt werden kann. Die vorangegangenen Abschnitte zeigen jedoch, dass im Stand der Technik und Forschung keine Vergleichsmethode existiert, mit der zuverlässig alle notwendigen oder nicht notwendigen Tests auf Gesamtfahrzeugebene klassifiziert werden können (vgl. Kap. 5.3). Die Bestimmung der „wahren“ Richtig-Positiven und der „wahren“ Richtig-Negativen ist daher für das Gesamtfahrzeug nicht möglich. Auch eine Ermittlung der Falsch-Positiven, d.h. der „zuviel“ gemachten Tests, scheitert, da die „wahre“ Menge der Richtig-Positiven nicht bekannt ist.

Die einzig validierbare Größe, die verbleibt, ist die Anzahl der falsch negativen Tests. Diese kann zwar nicht direkt ermittelt werden, weil die „wahre“ Menge der Richtig-Negativen unbekannt ist, sie kann aber aus Langzeitbeobachtungen von Entwicklungsprojekten abgeschätzt werden. Zu diesem Zweck können Daten über Fehler, die bei Produktbeobachtungen im Feld gesammelt werden, genutzt werden (vgl. Abschnitt 2.2). Beispiele für solche Informationsquellen sind nachträgliche Änderungen, d.h. Änderungen nach Serienanlauf, Werkstattfälle oder Rückrufaktionen. Für jeden Fehlerfall ist dann zu prüfen, ob er auf eine konkrete Änderung und die damit verbundene Testauswahl zurückzuführen ist.

Dazu ist eine detaillierte Analyse der Änderungshistorie des betroffenen Fahrzeugprojektes durchzuführen. Grundsätzlich sind dabei drei Analyseausgänge denkbar. Zunächst ist es möglich, dass der Test, der den aufgetretenen Fehler aufgedeckt hätte, gar nicht in der Testsuite T vorhanden ist. Dies weist auf eine Lücke im Testkatalog hin. Die Information kann genutzt werden, um die Testsuite zu vervollständigen. Aussagen über die Qualität der Testauswahlmethode lassen sich daraus jedoch nicht ableiten. Im zweiten Fall wurde der Test, der den Fehler aufgedeckt hätte, durchgeführt, der Fehler wurde allerdings nicht gefunden. Dies lässt darauf schließen, dass der Test falsch ausgeführt oder nicht korrekt ausgewertet wurde. Auch diese Information kann nicht zur Beurteilung der Güte der Testauswahl herangezogen werden. Sie kann allerdings verwendet werden, um die Testdurchführung oder –auswertung zu verbessern. Lediglich der dritte Fall lässt Rückschlüsse auf die Testauswahlmethode zu. Er betrifft die Situationen, in denen ein Test, der den Fehler aufgedeckt hätte, bei der Testauswahl nicht berücksichtigt wurde. Diese Beispiele können zur Bestimmung der Anzahl der Falsch-Negativen genutzt werden.

Aus einer langfristigen Bestimmung der Falsch-Negativen über mehrere Entwicklungsprojekte können dann statistische Aussagen über die Sicherheit der Testauswahl generiert werden. Damit ist zumindest eine teilweise Validierung der Güte der Testauswahl möglich. Nachteilig ist, dass nur der Aspekt der Sicherheit, nicht jedoch der Betrachtungswinkel der Fehlerrate i.S.v. falsch positiven Tests, evaluiert werden kann. Außerdem ist die geschilderte Validierungsmöglichkeit mit einem hohen Aufwand für die Datenerhebung und Aufbereitung verbunden.

8.3 Evaluierung der Effizienz der Testauswahl

Ein weiteres Beurteilungskriterium ist die Effizienz der Auswahlmethode. In der Literatur wird häufig das Einsparpotenzial im Vergleich zum Volltest zur Bewertung der Effizienz herangezogen.¹⁷⁰ Nachteilig daran ist, dass der Aufwand für die Testauswahl darin nicht berücksichtigt wird. In dieser Arbeit wird unter dem Begriff Effizienz daher der Nutzen der Testauswahl im Verhältnis zum Aufwand, der bei der Selektion der Testfälle entsteht, verstanden.

Nach Rothermel¹⁷¹ werden zwei Teilbereiche der Effizienz unterschieden. Zeitlich gesehen, ist eine Auswahlmethode effizient, wenn die Kosten für die Auswahl und Durchführung der Testfälle T' kleiner sind als die Kosten für die Durchführung der gesamten Testmatrix T :

$$K_{\text{Auswahl}}(T') + K_{\text{Durchführung}}(T') < K_{\text{Durchführung}}(T). \quad (8-7)$$

Daraus folgt:

$$K_{\text{Auswahl}}(T') < K_{\text{Durchführung}}(T - T'). \quad (8-8)$$

Der Aufwand für die Auswahl ist z.B. vom Automatisierungsgrad der Methode bzw. von der Implementierung in einem dafür eingesetzten Tool abhängig. Der zweite Teilbereich bewertet die technische Effizienz der Auswahlmethode. Dies beinhaltet den Umfang der Informationen, die für die Auswahlmethode benötigt werden. Dazu zählen z.B. Testhistorien oder zusätzliche Analyseinformationen, die erfasst, dokumentiert und gepflegt werden müssen. Außerdem fällt unter dieses Kriterium der Aufwand, der für die initiale Erstellung des Konzeptes notwendig ist. Eine Auswahlmethode sollte daher möglichst auf vorhandene Informationen und Tools zurückgreifen und in bestehende Prozesse integrierbar sein.

Um die praktische Anwendbarkeit der Methode sicherzustellen, ist eine hohe zeitliche Effizienz zwingend erforderlich. Im Folgenden wird daher der Fokus der Evaluierung

¹⁷⁰ Vgl. z.B. Nörenberg: Effizienter Regressionstest von E/E-Systemen, 2012, S. 187 ff.

¹⁷¹ Rothermel, Harrold: Analyzing Regression Test Selection Techniques, 1996

auf die zeitliche Effizienz gelegt. Ausführungen zur erwarteten technischen Effizienz sind in Abschnitt 9.3 zu finden.

Die Bestimmung der zeitlichen Effizienz gestaltet sich schwierig, da sie von vielen Faktoren abhängig ist. Faktoren, die die Effizienz beeinflussen sind nach Rothermel¹⁷² der Zeitpunkt der Anwendung der Methode im Produktlebenszyklus, der Zeitpunkt der Verfügbarkeit der benötigten Informationen, die Automatisierbarkeit sowie der Umfang der Systemevaluation. Rothermel schlägt daher eine empirische Bewertung der Kosten der Testauswahlmethode vor.

Leung und White¹⁷³ stellten ein detailliertes Modell zur Bewertung von Regressions-teststrategien auf. Sie beschränken sich auf die direkten Kosten, die mit den Tests selbst verbunden sind.¹⁷⁴ Darunter fallen die Personalkosten für den Tester sowie alle weiteren Ressourcen, die für die Testauswahl und -durchführung benötigt werden (z.B. Computer, Testlabor, Fahrzeuge, Teststrecke etc.). Leung und White unterscheiden ferner Systemanalysekosten, Testselektionskosten, Testdurchführungskosten und Ergebnisanalysekosten. Das Modell berücksichtigt allerdings nicht die Kosten, die entstehen, wenn ein Test ausgeschlossen wird, der eigentlich notwendig ist (z.B. Fehlerbehebung, Rückrufaktion etc.). Der Ansatz wurde daher von Malishevsky et al.¹⁷⁵ erweitert. Die Autoren beziehen die Fehlerkosten mit ein und berücksichtigen zusätzlich Wartungskosten, die durch die Methode verursacht werden.

Da das Modell von Malishevsky et al. insgesamt sehr komplex ist und viele einzelne Parameter enthält, die aufwändig ermittelt werden müssen, wird in dieser Arbeit ein vereinfachtes Kostenmodell vorgeschlagen, um die Effizienz der Testauswahlmethode zu evaluieren. Dazu werden für jeden einzelnen Test der Testsuite T drei Kostenfaktoren aufgestellt. Der erste Kostenfaktor K_D fasst die Aufwände zusammen, die bei der Durchführung des Tests entstehen. Dies sind z.B. die Personalkosten, die Kosten für das Fahrzeug oder die Rechnerkapazität und die Aufwände für die Bereitstellung der Teststrecke bzw. des Testlabors. Sie enthalten auch die Kosten für die Testauswertung. Die Testauswahlkosten K_A stellen die Aufwände dar, die im Rahmen der vorgestellten Testauswahlmethode anfallen, um einen Test bei einer Änderung ein- bzw. auszuschließen. Den dritten Kostenfaktor bilden die Fehlerkosten K_F . Sie beinhalten die Aufwände, die entstehen, wenn der Test fälschlicherweise nicht gemacht wird. Sie enthalten daher

¹⁷² Rothermel, Harrold: Analyzing Regression Test Selection Techniques, 1996

¹⁷³ Leung, White: A Cost Model to Compare Regression Test Strategies, 1991

¹⁷⁴ Vernachlässigte indirekte Kosten sind z.B. die Gemeinkosten für das allgemeine Testmanagement, die Kosten für die Datenspeicherung (u.a. Testfälle, Testergebnisse, Änderungshistorien) oder die Kosten für die Werkzeugentwicklung.

¹⁷⁵ Malishevsky et al.: Modeling the Cost-Benefits Tradeoffs for Regression Testing Techniques, 2002

die Kosten, die durch Fehler verursacht werden, die in diesem Test entdeckt werden können. Mit Hilfe dieser Kostenfaktoren kann eine Evaluierung der Effizienz der Testauswahlmethode durchgeführt werden. Die Testauswahl ist dann als effizient einzustufen, wenn die Durchführungskosten K_D , die Auswahlkosten K_A und die Fehlerkosten K_F für T' zusammen kleiner sind als die Durchführungskosten K_D von T .

$$K_D(T') + K_A(T') + K_F(T') < K_D(T) \quad (8-9)$$

Zur praktischen Umsetzung wird vorgeschlagen, die drei Kostenfaktoren für jeden Testfall in die Testbeschreibung mit aufzunehmen. Dabei besteht die Möglichkeit, die Kosten detailliert nach tatsächlichen Aufwänden (z.B. in €) anzugeben oder lediglich eine Klassifizierung in Aufwandsgruppen durchzuführen. Um statistisch zuverlässige Aussagen zu erhalten, ist es notwendig, die Erhebung dieser Daten langfristig, d.h. über mehrere Entwicklungsprojekte, durchzuführen.

8.4 Evaluierung der Generalität der Testauswahl

Auch die Generalität wird als Beurteilungskriterium für die Bewertung von Testauswahlmethoden herangezogen. Generalität bezeichnet nach Rothermel¹⁷⁶ die Fähigkeit einer Auswahltechnik, in einem weiten Bereich von Situationen anwendbar zu sein. Faktoren, die die Generalität beeinflussen, sind die Anwendbarkeit auf unterschiedliche Systeme, die Verwendbarkeit bezüglich verschiedener Arten von Änderungen, die Abhängigkeit von Annahmen und benötigten Informationen und die Verfügbarkeit von Tools.

Im Rahmen der Validierung der entwickelten Testauswahlmethode für den Umgang mit Änderungen in der Gesamtfahrzeugabsicherung ist daher zu überprüfen, ob die entwickelte Testauswahlmethode generell anwendbar ist. Dazu sind zwei Aspekte zu untersuchen.

Zum einen ist besonderer Fokus auf die Anwendbarkeit auf die Gesamtfahrzeugabsicherung zu legen. Diese besteht aus einer großen Anzahl an verschiedenen Tests mit jeweils unterschiedlichen Beurteilungszielen.¹⁷⁷ Um zu validieren, ob die Testauswahlmethode hier generell verwendbar ist, ist daher zu analysieren, ob die Methode für die verschie-

¹⁷⁶ Rothermel, Harrold: Analyzing Regression Test Selection Techniques, 1996

¹⁷⁷ Vgl. Kap. 2.4

denen Testkategorien in der Gesamtfahrzeugabsicherung anwendbar ist. Borgeest¹⁷⁸ nimmt eine Einteilung in acht unterschiedliche Kategorien von Tests vor. Dies sind Funktionstests, Robustheitstests, Benchmarks, Recovery-Tests, Konfigurations- und Kompatibilitätstests, Bedienbarkeitstests, Sicherheitstests und Lebensdauertests. Die Methode sollte auf Tests aller Kategorien anwendbar sein.

Zum anderen sollte bei der Validierung der Aspekt der Anwendbarkeit auf unterschiedliche Arten von Änderungen analysiert werden. Dazu ist zu untersuchen, ob sowohl Software- als auch Hardwareänderungen berücksichtigt werden. Es wird daher vorgeschlagen, die Generalität der entwickelten Testauswahlmethode anhand von exemplarischen Änderungen beider Änderungsarten in allen Testkategorien zu validieren.

Wenn die Methode für Tests aller Kategorien und für Änderungen auf Hardware- und Softwareebene einsetzbar ist, kann damit nachgewiesen werden, dass die Methode generell anwendbar ist. Falls sie für eine Test- oder Änderungskategorie nicht verwendet werden kann, d.h. falls die Auswahl von Testfällen in einzelnen Fällen scheitert, so kann eine Diskussion der Ursachen die Einschränkungen der Generalität offenlegen.

8.5 Fazit

Zur Validierung der vorgeschlagenen Testauswahlmethode werden insgesamt drei Untersuchungsziele definiert. Die Güte der Testauswahl bildet das erste Validierungsziel. Sie kann jedoch nur eingeschränkt geprüft werden, da keine Vergleichsmethode im Stand der Technik und Forschung existiert, die eine sichere Einteilung der Testmenge in „notwendig“ und „nicht notwendig“ ermöglicht. Es kann jedoch mithilfe von Langzeitdatenerfassungen eine statistische Aussage über die Sicherheit der Testauswahl gewonnen werden, die die falsch negativen Tests identifiziert. Die Ausführungen zeigen jedoch auch, dass für diese Validierung langfristige Datenerhebungen notwendig sind, so dass diese Analyse über einen großen Zeithorizont erfolgen muss. Das zweite Untersuchungsziel umfasst die Effizienz der Testauswahl. Dazu wird ein vereinfachtes Kostenmodell vorgeschlagen, das aus drei Faktoren besteht, die für jeden Testfall einzeln ermittelt werden müssen. Anhand dieser Daten ist ein Vergleich der Effizienz der Testauswahl zum Volltest möglich. Zur Durchführung fehlen allerdings auch hier noch Voraussetzungen, wie z.B. die Ermittlung der Aufwände pro Testfall. Zur Validierung der Generalität wird vorgeschlagen, die Methode anhand exemplarischer Änderungen auf acht unterschiedliche Testkategorien sowie zwei verschiedene Änderungsarten anzuwenden. Anhand einer Betrachtung dieser Änderungsbeispiele kann dann die Generalität überprüft werden.

¹⁷⁸ Borgeest: Elektronik in der Fahrzeugtechnik, 2010, S. 227 f.

Falls alle Untersuchungen positive Ergebnisse liefern, liegt eine sichere, effiziente und generell anwendbare Testauswahlmethode vor, die sich bewährt hat und der Nachweis der Realisierbarkeit ist erbracht. Sollte mindestens eine Untersuchung negative Ergebnisse liefern, kann eine Diskussion der Ursachen und somit ein Rückschluss auf die Realisierbarkeit der Testauswahlmethode erfolgen. Das Validierungskonzept ermöglicht damit eine Aussage darüber, ob die Methode ihrem Einsatzzweck, der Überprüfung von Änderungen in der Gesamtfahrzeugabsicherung, gerecht werden kann.

9 Schlussfolgerungen für den Entwicklungsprozess

Aus dem vorgeschlagenen Gesamtkonzept für die Testauswahl ergeben sich Schlussfolgerungen für den zugrundeliegenden Entwicklungsprozess. Diese werden im folgenden Kapitel dargelegt. Zunächst erfolgt eine Beschreibung der Anforderungen an die Entwicklung von Softwarekomponenten. Dies wird unter dem Begriff „Design for Exclusion“ zusammengefasst. Darauf aufbauend werden die Anforderungen an das Testdesign diskutiert. Das Kapitel endet mit einem Fazit.

9.1 Design for Exclusion

Aus den Konkretisierungen der Ausschlussbeweise leiten sich Anforderungen an die Entwicklung von Softwarekomponenten ab. Dafür wird der Begriff „Design for Exclusion (DfE)“ eingeführt. Das DfE bildet ein Rahmenwerk für die Funktionsentwicklung, das die exklusionsbasierte Testauswahlmethode unterstützt. Der erste Ausschlussbeweis (AB I: Softwareänderungen) (siehe 6.2.1) nutzt die Definition von Arbeitsbereichen von Funktionen. Eine Bedingung, um diesen Exklusionsbeweis durchführen zu können, ist damit die Definition dieser Arbeitsbereiche in der Software. Ein Arbeitsbereich einer Funktion markiert dabei die erklärten Voraussetzungen, unter denen diese Funktion aktiv ist. Hier werden somit die Situationsparameter festgelegt, unter denen die Funktion arbeitet. Folgende Kriterien sind zur Definition der Arbeitsbereiche heranzuziehen:

Explizität (Festforderung)

Der Arbeitsbereich ist explizit in der Software abzulegen (z.B. Aktivierung nur bei Querbeschleunigung $> 3 \text{ m/s}^2$). Dies ist nötig, um die Nachweisführung der Einhaltung des Arbeitsbereichs effizient durchzuführen (siehe 6.2.1). Das Konzept ist in der Literatur unter dem Begriff Zusicherungen („assertions“) bekannt.¹⁷⁹ Die konkrete softwaretechnische Umsetzung ist noch zu klären. Da die Implementierung jedoch ein hohes Expertenwissen erfordert, ist vermutlich eine manuelle Integration notwendig.

¹⁷⁹ Liggesmeyer: Software-Qualität, 2009, S. 211 ff.

Korrespondenz (Festforderung)

Um den Vergleich der Arbeitsbereiche der Funktionen mit den Einsatzgrenzen der Tests auszuführen, sind die Arbeitsbereiche der Funktionen an die Testeinsatzgrenzen anzulehnen. Da an dieser Stelle Tests in der Gesamtfahrzeugabsicherung betrachtet werden, sollten für die Definition der Arbeitsbereiche daher die Parameter herangezogen werden, die auch zur Beschreibung der Fahrsituationen in der Gesamtfahrzeugabsicherung verwendet werden.

Unzweifelhaftigkeit (Festforderung)

Um sicherzustellen, dass die definierten Arbeitsbereiche der Funktionen im Fahrbetrieb auch eingehalten werden, müssen die Bedingungen, die den Arbeitsbereich festlegen, vom System unzweifelhaft erkannt werden. Die Unzweifelhaftigkeit ist dabei für jedes Signal und jeden Testfall zu definieren (z.B. Eigendiagnoseschwelle). Daraus resultiert, dass einfache Signale zur Definition der Arbeitsbereiche besser geeignet sind als z.B. Parameter, die in mehreren Berechnungsschritten und unter vereinfachenden Modellannahmen berechnet werden.

Detaillierungstiefe (Zielforderung)

Je mehr Gemeinsamkeiten zwischen Testeinsatzgrenzen und Arbeitsbereichen definiert sind, desto mehr Exklusionsmöglichkeiten bestehen. Es ist daher zielführend, die Arbeitsbereiche der Funktionen ebenso wie die Einsatzgrenzen der Tests möglichst detailliert zu beschreiben, um die Anzahl der durchgeführten Tests bei Änderungen durch den Ausschlussbeweis zu maximieren.

Allgemeingültigkeit (Zielforderung)

Allgemeine Arbeitsbereiche (d.h. unabhängig vom Fahrzeugprojekt) erleichtern die Nachvollziehbarkeit und Wiederverwendbarkeit der Ausschlussbeweise. Es wird daher empfohlen, die Arbeitsbereiche allgemeingültig zu formulieren.

Neben den Anforderungen an die Definition von Arbeitsbereichen existieren in der Softwaretechnik viele Methoden, die den Einsatz der exklusionsbasierten Testauswahlmethode unterstützen. Starke¹⁸⁰ schlägt z.B. vor, durch die Minimierung von Abhängigkeiten in der Software, die Flexibilität der Architektur für zukünftige Änderungen zu erhöhen. Dazu werden in den Entwürfen die Abhängigkeiten zwischen Komponenten derart konstruiert, dass künftige Änderungen keine neuen Abhängigkeiten erzeugen. So können die Auswirkungen von Änderungen minimiert und die Wahrscheinlichkeit von Exklusionsmöglichkeiten erhöht werden. Starke stellte diverse Entwurfsprinzipien vor,

¹⁸⁰ Starke: Effektive Softwarearchitekturen, 2011, S. 154 ff.

die bei der Entwicklung zu berücksichtigen sind (z.B. lose Kopplung, hohe Kohäsion, Offen-Geschlossen-Prinzip, Abhängigkeit nur von Abstraktionen, Abtrennung von Schnittstellen, zyklische Abhängigkeiten vermeiden, Liskov-Substitutionsprinzip).

Ein anderes Konzept, das die Exklusionsmöglichkeiten erhöhen könnte, ist das „Design for Changeability - DfC“ nach Schulz¹⁸¹ und Fricke¹⁸². Der Ansatz beruht darauf, die Änderbarkeit einer Systemarchitektur über den gesamten Lebenszyklus zu berücksichtigen. Die Strategie zielt darauf ab, späte Änderungen zu ermöglichen, um damit jeden Entwicklungsschritt eines Systems kosteneffizient zu gestalten. So sollen Unternehmen befähigt werden, auf aufkommende und veränderte Kundenanforderungen und Märkte zu reagieren.¹⁸³ Schulz et al.¹⁸⁴ nennen vier Aspekte, die die Fähigkeit eines Systems, mit Änderungen von innen oder aus der Umgebung umzugehen, ausbilden. Dies sind Agilität, Adaptivität, Flexibilität und Robustheit. Um diese Eigenschaften in einem System zu implementieren, werden verschiedene Gestaltungsprinzipien für die Produktentwicklung vorgeschlagen. Dies sind z.B. Idealität / Einfachheit, Unabhängigkeit, Modularität / Kapselung, Integrationsfähigkeit, Dezentralisierung, Skalierbarkeit / Selbstähnlichkeit und nicht-hierarchische Integration.¹⁸⁵

Auch Ring¹⁸⁶ stellt Gestaltungsprinzipien zum effizienten Umgang mit Änderungen vor, sog. „Agile System Design Principles“. Diese umfassen z.B. abgeschlossene Module / Kapseln, Austauschbarkeit und Redundanz.

Welche der genannten Ansätze im Bezug auf das DfE vielversprechend anwendbar sind, ist noch weiter zu untersuchen.

9.2 Anforderungen an das Testdesign

Aus den Umsetzungsvorschlägen für die exklusionsbasierte Testauswahlmethode ergeben sich außerdem Anforderungen an das Testdesign, die nachfolgend dargelegt werden. Sie bilden ein Rahmenwerk für die Testentwicklung, das die Exklusionsbeweissführung unterstützt.

¹⁸¹ Schulz, Fricke: Incorporating Flexibility, Agility, Robustness, and Adaptability, 1999

¹⁸² Fricke: Der Änderungsprozess als Grundlage einer nutzerzentrierten Systementwicklung, 1999

¹⁸³ Schulz, Fricke: Incorporating Flexibility, Agility, Robustness, and Adaptability, 1999

¹⁸⁴ Schulz et al.: Enabling Changes in Systems, 2000; Schulz, Fricke: Incorporating Flexibility, Agility, Robustness, and Adaptability, 1999

¹⁸⁵ Schulz et al.: Enabling Changes in Systems, 2000

¹⁸⁶ Ring: Rapid Evolution of all your Systems, 1998

Definition von Testeinsatzgrenzen (Festforderung)

Der erste Ausschlussbeweis (AB I: Softwareänderung) (siehe Abschnitt 6.2.1) führt einen Vergleich der Einsatzgrenzen der Tests mit dem Arbeitsbereich der geänderten Funktion durch. Um diesen Vergleich durchführen zu können, sind die Situationsparameter der Fahrmanöver **explizit** zu beschreiben. Für die Definition der Testeinsatzgrenzen gelten die gleichen Anforderungen wie für die Festlegung der Arbeitsbereiche der Funktionen: Explizität, Korrespondenz, Unzweifelhaftigkeit, Detaillierungstiefe und Allgemeingültigkeit (siehe Abschnitt 9.1).

Definition von physikalischen Auswirkungsklassen (Festforderung)

Um den ersten Ausschlussbeweis für Hardwareänderungen (siehe Abschnitt 6.2.2) anwenden zu können, sind den Testfällen die physikalischen Auswirkungsklassen **explizit** zuzuordnen. Es wird vorgeschlagen, die Klassifikation von Hirtz et al.¹⁸⁷ zu nutzen. Da in dem untersuchten Anwendungsbereich der Gesamtfahrzeugabsicherung vor allem die Energieflüsse von Bedeutung sind, bieten sich diese für eine erste Implementierung an. Hirtz et al. schlagen zwölf Kategorien für Energieflüsse vor: menschlich, akustisch, biologisch, chemisch, elektrisch, elektromagnetisch, hydraulisch, magnetisch, mechanisch, pneumatisch, radioaktiv/nuklear und thermisch. Um den Ausschlussbeweis durchführen zu können, ist es erforderlich, die Auswirkungskategorien **vollständig** zu definieren, d.h. auch die impliziten Auswirkungen sind für den jeweiligen Test zu definieren. Ansonsten besteht die Gefahr, dass Tests ausgeschlossen werden, die zur Absicherung der Änderung erforderlich sind (falsch negative Tests).

Definition von Testzielen (Festforderung)

Der zweite Ausschlussbeweis (AB II) (siehe Abschnitt 6.2.3) führt eine Testzieloptimierung durch. Voraussetzung für die Nachweisführung ist eine **explizite** Definition der Ziele der Tests. Jedes Testziel beinhaltet dabei die zu untersuchende Eigenschaft (z.B. Fahrkomfort, Bremsweg, Agilität) inklusive einer Beschreibung der Fahrsituation, in der diese Eigenschaft bewertet wird (z.B. bei Maximalgeschwindigkeit, auf Hochreihwert, im 3. Gang). Es ist dabei erforderlich, die Ziele der einzelnen Tests **vollständig** zu beschreiben. Sonst besteht auch hier das Risiko der Vernachlässigung von notwendigen Tests (f_n).

Definition von Kostenfaktoren (Festforderung)

Um eine Evaluierung der Effizienz der Testauswahl durchführen zu können, ist es zudem erforderlich, jedem Testfall die drei Kostenfaktoren K_D , K_A und K_F zuzuordnen. Sie beinhalten die Aufwände, die bei der Durchführung, Auswahl und im Fehlerfall (f_n) entstehen (siehe Abschnitt 8.3).

¹⁸⁷ Hirtz et al.: A functional basis for engineering design, 2002

Toolunterstützung (Zielforderung)

Um die Ausschlussbeweissführung möglichst effizient zu gestalten, wird eine Verwaltung der Testdaten in einem Tool (z.B. in einer Datenbank) vorgeschlagen. Damit werden Filterungen (z.B. nach Testeinsatzgrenzen, Auswirkungsklassen oder Testzielen) erleichtert.

9.3 Fazit und Diskussion

Die beiden vorangegangenen Abschnitte zeigen, dass die Implementierung des vorgeschlagenen Testauswahlkonzeptes viele Anpassungen des heutigen Entwicklungsprozesses erfordert. Der hohe initiale Aufwand sowie der enorme Umfang an Informationen, die erfasst, dokumentiert und gepflegt werden müssen, weisen darauf hin, dass der vorgestellte Ansatz zu Beginn eine geringe technische Effizienz (vgl. Abschnitt 8.3) aufzeigt. Da es sich jedoch um eine deduktiv abgeleitete Methode handelt, sind die geforderten Änderungen als zwingende Notwendigkeit für eine sichere Testauswahl anzusehen. Erst wenn die beschriebenen Anpassungen im Entwicklungsprozess durchgeführt werden, ergibt sich durch die vorgestellte Testauswahlmethode eine Chance auf eine günstige Nutzen-Kosten-Relation und auch erst dann kann sie validiert werden.

Die Implementierung der Maßnahmen ist jedoch mit hohem Aufwand verbunden. So ist z.B. für den ersten Ausschlussbeweis (AB I: Softwareänderungen) (siehe Abschnitt 6.2.1) die Definition von Arbeitsbereichen der Funktionen notwendig. In der Softwareentwicklung ist mithilfe der Modulspezifikation formal sicherzustellen, dass dieser Arbeitsbereich immer eingehalten wird. Nur dann kann der Exklusionsbeweis für die Testauswahl verwendet werden. Um die genannten Anforderungen zu implementieren, sind umfangreiche Änderungen in den einzelnen Funktionsmodulen erforderlich. Es stellt sich daher die Frage, ob ein Festhalten an den bisher gewachsenen Systemarchitekturen langfristig noch erfolversprechend ist oder ob nicht vielmehr neue Systemstrukturen zu entwerfen sind, die die Änderbarkeit und Testbarkeit stärker in den Fokus rücken. Dies ist vor allem vor dem Hintergrund der Tatsache zu betrachten, dass die Änderungshäufigkeit der gegenwärtigen, stark vernetzten Systeme stark zunimmt.¹⁸⁸ Dies ist darauf zurückzuführen, dass der heutige Wettbewerb zu einem großen Teil über Innovationen stattfindet und diese sich zu 90 % in der Elektronik und Software des Fahrzeugs widerspiegeln.

Auch im Bereich des Testmanagements sind die nötigen Änderungen umfangreich. So wird z.B. eine vollständige Beschreibung aller Testziele jedes Testfalls benötigt. Dies

¹⁸⁸ Broy et al.: Architekturen softwarebasierter Funktionen im Fahrzeug, 2011, S. 49

erfordert eine Überarbeitung der kompletten Testsuite der Gesamtfahrzeugabsicherung, bei der für viele Tests keine konkreten Testziele formuliert sind, sondern vielmehr lediglich die Angabe einer km-Angabe erfolgt (z.B. 100.000 km Dauerlauferprobung). Es stellt sich daher die Frage, ob diese Investitionen den Nutzen überhaupt rechtfertigen. Dabei ist jedoch zu beachten, dass die Testziele nicht nur zur Anwendung der gezeigten Testauswahlmethode verwendet werden können, sie können auch der Verbesserung heutiger Spezifikationen dienen. So können Testziele, für die heute keine Anforderungen existieren, in die Spezifikation aufgenommen werden, wodurch die Vollständigkeit der Anforderungen verbessert wird. Zudem kann durch die Überarbeitung sichtbar werden, an welchen Stellen Anforderungen nicht präzise genug formuliert sind, so dass die Anforderungen systematisch überprüft werden. Mithilfe der vollständigen Testziele kann durch die Implementierung der vorgeschlagenen Methode damit letztlich ein Migrationspfad von der feldbasierten zur spezifikationsbasierten Absicherung (vgl. Abschnitt 2.5.1) begründet werden.

Durch die Definition von Testeinsatzgrenzen und physikalischen Auswirkungsklassen wird es zudem möglich, die Gesamtfahrzeugabsicherung umfassend zu analysieren. So kann z.B. die Testabdeckung in unterschiedlichen Fahrsituationen bestimmt werden. Damit kann dann eine Bewertung der Qualität des empirisch ermittelten Testkollektivs erfolgen. Außerdem können die Informationen genutzt werden, um Redundanzen im Testkatalog zu identifizieren, wodurch die Gesamtfahrzeugabsicherung insgesamt optimiert werden kann.

Die Aufstellung von detaillierten Kosten je Testfall kann darüber hinaus genutzt werden, Aufwandstreiber (Tests mit hohen Durchführungskosten K_D) und Risikotests (hohe Fehlerkosten K_F) zu ermitteln. Aufwandstreiber bieten ein hohes Potenzial für Kostenreduktionen, wenn ein anderes Testverfahren gefunden werden kann, das die gleichen Testziele bei weniger Aufwand ermittelt (z.B. SiL oder HiL-Test anstatt eines Tests im Realfahrzeug). Sie geben damit Hinweise auf Verbesserungsmöglichkeiten. Risikotests können zur Priorisierung von Testfällen unter gegebenen Randbedingungen (z.B. zur Verfügung stehender Zeit) verwendet werden.

Die vorgeschlagenen Maßnahmen führen damit dazu, dass das Testmanagement in der Gesamtfahrzeugabsicherung insgesamt transparenter wird. Ferner wird das Know-How, das heute zum großen Teil im Expertenwissen der zuständigen Mitarbeiter liegt, festgehalten, wodurch es breiter nutzbar und leichter übertragbar wird und auch wieder in den Entwicklungsprozess zurückgeführt werden kann.

Die beschriebenen Ausführungen zeigen, dass für die Implementierung der Testauswahlmethode noch Voraussetzungen fehlen. Mit der Schaffung dieser Voraussetzungen lässt sich jedoch auch die Prozessqualität in der Gesamtfahrzeugabsicherung verbessern. Dennoch bleibt die Frage ungeklärt, ob der hohe Aufwand für die Implementierung der Testauswahlmethode ihren Nutzen rechtfertigt. Eine Alternative könnte die Umstellung

der Gesamtfahrzeugabsicherung auf die spezifikationsbasierte Absicherung sein, die eine einfache Auswahl der Testfälle bei Änderungen ermöglicht (siehe Abschnitt 2.5.3). Doch auch dies ist mit vielen notwendigen Änderungen im Entwicklungsprozess verbunden. Teilweise sind dafür sogar dieselben Maßnahmen (z.B. vollständige Testziele) erforderlich. Das vorgeschlagene Testauswahlkonzept kann daher sogar dabei unterstützen, mithilfe seiner notwendigen Anpassungen (Design for Exclusion, Testmanagement) einen Beitrag in Richtung des durchgängigen Einsatzes der spezifikationsbasierten Absicherung zu leisten.

10 Zusammenfassung und Ausblick

Die wachsende Anzahl von Softwarefunktionen sowie die erhöhte Varianten- und Ausstattungsvielfalt heutiger Fahrzeuge führen zu einer gesteigerten Komplexität automobiler Systeme. Begleitet wird dies von stetig wachsenden Anforderungen an die Qualität und Zuverlässigkeit der Produkte. Zusammengenommen folgt daraus ein enormer Zuwachs der Bedeutung der Gesamtfahrzeugabsicherung, die die letzte Teststufe des Entwicklungsprozesses vor Auslieferung des Produkts an den Kunden darstellt.

Die einleitende Analyse der Rahmenbedingungen liefert einen Beitrag zur systematischen Untersuchung der Potenziale und Defizite der Gesamtfahrzeugabsicherung im Entwicklungsprozess in der Automobilindustrie. Dazu wird eine umfassende Modellvorstellung für den Absicherungsprozess entwickelt. Darüber hinaus führt die Analyse zur Differenzierung von zwei Paradigmen der Gesamtfahrzeugabsicherung. Das erste Paradigma stellt die spezifikationsbasierte Absicherung dar. Sie beruht auf einer Testfallgenerierung, die auf einer lückenlosen und eindeutigen Spezifikation aufbaut. Von Vorteil bei diesem Konzept ist, dass der Aufwand für die Testfallgenerierung selbst niedrig ausfällt. Außerdem liefern die Anforderungen klare Kriterien für die anschließende Freigabe, so dass diese rechtlich relevante Entscheidung objektivierbar wird. Nachteilig ist, dass die Qualität der Absicherung stark von der Qualität der Spezifikation abhängt. Die Umsetzung der spezifikationsbasierten Absicherung wird heute angestrebt (u.a. durch die Anforderungen der ISO 26262), sie ist aber nicht Stand der Technik, da noch Voraussetzungen für ihre Anwendung fehlen. Das zweite Paradigma, die feldbasierte Absicherung, stützt sich auf eine Abbildung der Nutzung des Fahrzeugs im Betrieb (sog. Prognosewerkzeug). Dazu wird die Absicherung derart durchgeführt, dass abhängig vom Nutzungsprofil des Fahrzeugs ein repräsentatives Kollektiv von Fahrsituationen überprüft wird. In der Praxis dominiert dabei die empirische Testfallgenerierung. Vorteil dieses Vorgehens ist, dass auf bestehende Testfälle zurückgegriffen werden kann, so dass der Aufwand für die Testfallgenerierung gering ausfällt. Zur Bestimmung der Qualität existieren derzeit jedoch keine Bewertungsmaßstäbe. Dies führt insbesondere bei der Freigabe zu Schwierigkeiten, da keine eindeutigen Kriterien ermittelt werden können, nach denen freigegeben wird. Der Vergleich zeigt, dass Testkonzepte, die allein auf dem Paradigma der spezifikationsbasierten Absicherung aufbauen, aufgrund der noch fehlenden Voraussetzungen derzeit nicht ausreichend sind. Die feldbasierte Absicherung ist daher trotz ihrer Nachteile unverzichtbar.

Eine weitere Erkenntnis aus der Analyse der Gesamtfahrzeugabsicherung besteht in der Identifizierung von fünf zentralen Herausforderungen. Sie zeigen Rahmenbedingungen auf, die bei Veränderungen der Gesamtfahrzeugabsicherung zu berücksichtigen sind.

Die hohe rechtliche Relevanz der Gesamtfahrzeugabsicherung kennzeichnet die erste Herausforderung. Die zweite Schwierigkeit stellt die hohe Systemkomplexität des Fahrzeugesamtsystems dar. Sie führt dazu, dass ein vollständiges Testen einer Funktion unter ökonomisch vertretbarem Aufwand nicht mehr realisierbar ist. Der hohe Aufwand für die Durchführung der Absicherungsprozesse ist die dritte Herausforderung. Die vierte Schwierigkeit stellt der späte Zeitpunkt der Gesamtfahrzeugabsicherung im Entwicklungsprozess dar. Änderungen betreffen die fünfte Problemstellung heutiger Absicherungsprozesse. Sie erfordern eine (teilweise) Wiederholung der Absicherungstests, so dass die begrenzt zur Verfügung stehenden Ressourcen stärker belastet werden.

Der letzten Herausforderung widmet sich der Hauptteil dieser Arbeit. Das Änderungsmanagement ist in der Automobilindustrie ein entscheidender Wettbewerbsfaktor. Um Änderungen umsetzen zu können, ist eine erneute Absicherung des Fahrzeugs erforderlich, bei der die Auswirkungen der Änderung mit Hilfe von Tests überprüft werden. Die Auswahl der Testfälle erfolgt überwiegend expertenbasiert und manuell. Die Qualität des Ergebnisses ist damit stark von der Erfahrung des Testers abhängig. Eine Hilfestellung in Form einer nachvollziehbaren Methodik, die vor allem auch eine rechtliche Absicherung ermöglicht, existiert derzeit nicht. In der Praxis erfolgt daher häufig die komplette Wiederholung aller Absicherungstests, was einen enormen Erprobungsaufwand mit sich bringt.

Diese Arbeit untersucht daher anhand einer deduktiven Vorgehensweise die Fragestellung, ob eine Methode entwickelt werden kann, die eine sichere, nachvollziehbare systematische Auswahl von Testfällen bei Änderungen ermöglicht. Die Methode sollte auf die Gesamtfahrzeugabsicherung anwendbar sein und den Aufwand für die Absicherung im Vergleich zum Volltest verkleinern. Dabei ist zur Reduktion von Produkthaftungsrisiken vor allem die Sicherheit der nachfolgenden Freigabeentscheidung zu verbessern.

Die Abstraktion des Problems der Testauswahl zur Absicherung von Änderungen auf Gesamtfahrzeugebene identifiziert zwei prinzipielle Lösungsmöglichkeiten. Der Inklusionsansatz bestimmt aus der Menge aller Tests die notwendigen Testfälle und überführt sie in die Menge der durchzuführenden Tests. Der Exklusionsansatz hingegen ermittelt die nicht notwendigen Testfälle und transferiert sie in die Menge der nicht durchzuführenden Tests. Ein grundsätzlicher Vergleich der beiden Lösungskonzepte anhand der in der Arbeit entwickelten Kriterien Selektionsqualität, Selektionskosten, Selektionsaufwand sowie Umgang mit mehreren Änderungen zeigt, dass unter der Annahme einer idealen Selektionsqualität auf dieser Abstraktionsebene beide Konzepte als gleichwertig einzustufen sind.

Die weitere Konkretisierung legt dar, dass für die inklusionsbasierte Testauswahl nach dem Stand der Technik und Forschung keine effektive Methode existiert, mit der alle fehleraufdeckenden Tests ermittelt werden können. Es ist lediglich möglich, eine Über-

menge von Testfällen zu bestimmen, die von der Änderung betroffen sind. Dabei können jedoch falsch positive Testfälle nicht ausgeschlossen werden. Vorteil vieler in der Literatur beschriebener und in der Praxis eingesetzter Verfahren der inklusionsbasierten Testauswahl ist, dass sie sehr effizient einsetzbar sind. Die zugrundeliegenden Modelle für die Testauswahl sind teilweise schon vorhanden und auch die Testauswahl ist automatisiert durchführbar, so dass der Selektionsaufwand als gering eingestuft werden kann. Ein großer Nachteil des Inklusionskonzeptes ist jedoch, dass die Qualität der Testauswahl von der Qualität des zugrundeliegenden Modells zur Ermittlung der Wechselwirkungen der Änderung abhängig ist. Um eine sichere Testauswahl durchführen zu können, muss nachgewiesen werden, dass das Modell ein vollständiges und korrektes Abbild des Gesamtfahrzeugs darstellt. Da die Menge der Wechselwirkungen unbekannt ist und absolute Aussagen über eine unbestimmte oder unendlich große Menge grundsätzlich nicht möglich sind, kann dieser Beweis lediglich induktiv geführt werden. Es ist daher offen, wie die inklusionsbasierte Testauswahl auf Gesamtfahrzeugebene jemals beweisen kann, dass sie diese Voraussetzungen erfüllt. Dies führt zu der Schlussfolgerung, dass - obwohl in der Literatur sehr viele Inklusionsansätze beschrieben sind - nach derzeitigem Stand der Technik und Forschung keine sichere Auswahl von Testfällen in der Gesamtfahrzeugabsicherung mithilfe des Inklusionsprinzips möglich ist.

Die Untersuchungen zeigen, dass in der Literatur keine beispielhaften Konkretisierungen des Exklusionsprinzips auf Gesamtfahrzeugebene existieren. In dieser Arbeit wird daher erstmalig eine exemplarische Umsetzung dieses Konzeptes anhand von zwei Ausschlussbeweisen entwickelt, die anhand von drei Ausprägungsbeispielen umgesetzt werden. Die entworfene Realisierung des Exklusionsansatzes gewährleistet eine hohe Vollständigkeit der Testauswahl, so dass die resultierende Testmenge immer sicher (i.S.v. keine falsch negativen Tests) ist. Dies macht eine Anwendung im rechtlich relevanten Bereich der Gesamtfahrzeugabsicherung möglich. Durch die Exklusion von Falsch-Positiven wird zudem die Fehlerrate reduziert. Eine Fehlerrate von null kann allerdings nur erreicht werden, wenn alle Tests, die nicht notwendig sind, exkludiert werden. Wie die exemplarische Umsetzung zeigt, ist dieser Ausschlussbeweis in der Praxis jedoch sehr aufwändig, da ein Testfall nur ausgeschlossen werden kann, wenn er nachweislich in keiner Wechselwirkung mit der Änderung steht. Daraus folgt, dass die Methode mit einem sehr hohen Selektionsaufwand verbunden ist. Jede Unsicherheit in der Beurteilung führt darüber hinaus automatisch zu einer Inklusion des Tests. Dies hat zur Folge, dass in der Menge der durchgeführten Tests auch Testfälle enthalten sein können, die eigentlich nicht notwendig sind, für die jedoch der Beweis der Wechselwirkungsfreiheit nicht möglich ist (falsch positive Tests).

Das in dieser Arbeit anhand einer Effizienzbetrachtung entwickelte Gesamtkonzept vereint die spezifischen Vorteile beider prinzipieller Lösungsmöglichkeiten. Ausgangsbasis für die vorgeschlagene zweistufige Testauswahl ist die Menge aller Tests, die zur Absicherung des Gesamtsystems herangezogen werden. Im ersten Auswahlschritt wird

die Inklusionsanalyse durchgeführt. Dazu eignen sich besonders die Methoden aus dem Stand der Technik, die sich in heutige Prozesse einfach integrieren lassen. So identifiziert die Inklusionsanalyse alle Tests, bei denen eine Wechselwirkung mit der Änderung wahrscheinlich ist. Die Testfälle, bei denen die Inklusionsanalyse keine Hinweise auf Wechselwirkungen ermittelt hat, bieten somit ein hohes Potenzial für den Ausschluss bei vertretbarem Aufwand. Sie bilden die Ausgangsbasis für die Exklusionsanalyse, die den zweiten Auswahlschritt des Testkonzeptes darstellt. Hier wird eine Exklusion von Testfällen durchgeführt, bei denen Wechselwirkungen der Änderung nachweislich nicht vorhanden sind. Tests, für die dieser Ausschluss nicht möglich ist, werden durchgeführt. Das vorgeschlagene Gesamtkonzept ermöglicht eine sichere und effiziente Identifizierung von Testfällen, die bei der Überprüfung von Änderungen in der Gesamtfahrzeugabsicherung entfallen können. Eine Einschränkung des Konzeptes ergibt sich lediglich in den Fällen, in denen die Änderung eine Funktionserweiterung darstellt, die einen Testfall erfordert, der nicht im Ausgangstestkatalog enthalten ist.

Der beschriebene Validierungsansatz für das aufgestellte Testauswahlkonzept definiert drei Untersuchungsziele. Zur Evaluierung der Güte des Ansatzes können mithilfe von Langzeitdatenerfassungen statistische Aussagen über die Sicherheit der Testauswahl gewonnen werden, indem die Anzahl der falsch negativen Tests ermittelt wird. Die Effizienz des Konzeptes kann anhand eines vereinfachten Kostenmodells beurteilt werden. Zur Validierung der Generalität wird vorgeschlagen, die Testauswahlmethode anhand von exemplarischen Änderungsbeispielen auf acht unterschiedliche Testkategorien sowie zwei verschiedene Änderungsarten anzuwenden. Somit liegt ein umfassendes Validierungskonzept für die Testauswahlmethode vor, für dessen Anwendung allerdings noch Voraussetzungen fehlen.

Eine Diskussion der Schlussfolgerungen für den Entwicklungsprozess identifiziert notwendige Anforderungen der in dieser Arbeit aufgestellten Methode an die Softwareentwicklung sowie das Testdesign. Erst wenn die beschriebenen Maßnahmen durchgeführt werden, ergibt sich durch die vorgestellte Testauswahlmethode ein Nutzen und auch erst dann kann sie validiert werden. Die Umsetzung der Anpassungen ist jedoch mit hohem Aufwand verbunden. So sind umfangreiche Änderungen in den einzelnen Funktionsmodulen der Software erforderlich. Außerdem wird eine Überarbeitung der gesamten Testsuite der Gesamtfahrzeugabsicherung notwendig. Die Untersuchung zeigt jedoch, dass mit der Schaffung dieser Voraussetzungen auch die Prozessqualität in der Gesamtfahrzeugabsicherung verbessert werden kann. Dennoch bleibt die Frage ungeklärt, ob der hohe Aufwand für die Implementierung der Testauswahlmethode ihren Nutzen rechtfertigt. Als Alternative wird die Umstellung der Gesamtfahrzeugabsicherung auf die spezifikationsbasierte Absicherung diskutiert, die eine einfache Auswahl der Testfälle bei Änderungen ermöglicht. Doch auch dies ist mit vielen notwendigen Änderungen im Entwicklungsprozess verbunden. Teilweise sind dafür sogar dieselben Maßnahmen (z.B. vollständige Testziele) erforderlich. Dies führt zu der Schlussfolge-

zung, dass das vorgeschlagene Testauswahlkonzept mithilfe seiner notwendigen Anpassungen einen Beitrag in Richtung des durchgängigen Einsatzes der spezifikationsbasierten Absicherung leisten kann.

Zusammenfassend lässt sich damit feststellen, dass mit der in dieser Arbeit entwickelten Testauswahlmethode ein deduktiv abgeleitetes Konzept vorliegt, das durch die Anwendung des Exklusionsprinzips erstmalig eine sichere Identifizierung von Testfällen bei Änderungen in der Gesamtfahrzeugabsicherung ermöglicht. Dies ist insbesondere vor dem Hintergrund der rechtlichen Relevanz dieses Prozesses ein nicht zu unterschätzender Vorteil. Zudem kombiniert der Ansatz die spezifischen Vorteile des Inklusions- und Exklusionsprinzips. Dadurch wird sichergestellt, dass die Testauswahlmethode effizient einsetzbar ist. Ob damit auch eine Reduktion der Testfälle im Vergleich zur heutigen expertenbasierten Testfallauswahl erzielt werden kann, bleibt offen. Der hohe Initialaufwand, der mit der Implementierung des Ansatzes verbunden ist, kann dadurch gerechtfertigt werden, dass das Konzept die Prozessqualität im Bereich Software- und Testdesign nachhaltig verbessert und einen Migrationspfad von der feldbasierten zur spezifikationsbasierten Absicherung darstellt.

Um die entwickelte Methode in Absicherungsprozessen in der Automobilindustrie anwenden zu können, sind im ersten Schritt die in dieser Arbeit beschriebenen Veränderungen in der Softwareentwicklung sowie im Testmanagement umzusetzen. Dies kann z.B. als Teil einer Qualitätsverbesserungsoffensive geschehen. Auf dieser Grundlage kann dann eine Überprüfung der Realisierbarkeit der Testauswahlmethode gemäß dem vorgeschlagenen Validierungskonzept erfolgen, so dass der Nutzen und die Einschränkungen der Methode abgeschätzt werden können.

In Zukunft ist es zudem erforderlich, den ersten Ausschlussbeweis für Softwareänderungen (AB I) weiter auszuarbeiten. Grundlage für den Ausschlussbeweis ist ein Nachweis der Einhaltung des Arbeitsbereichs einer Funktion. Wie die Umsetzung dieses Nachweises erfolgen kann, ist noch offen. Eine mögliche Lösung bilden statische Verifikationsverfahren (z.B. formaler Korrektheitsbeweis oder symbolischer Test), die es gestatten, vollständige Aussagen über Softwareeigenschaften zu erhalten. Für eine mögliche Anwendung ist zu prüfen, ob die erforderlichen Voraussetzungen durch die vorliegende Automobilsoftware abgedeckt werden bzw. welche Änderungen damit in heutigen Softwarekomponenten notwendig sind.

Forschungsbedarf besteht auch in der weiteren Konkretisierung des Rahmenwerks für die Funktionsentwicklung (Design for Exclusion). Dabei ist zu untersuchen, welche Methoden aus der Softwaretechnik, den Einsatz der exklusionsbasierten Testauswahl weiter unterstützen können. Gestaltungsprinzipien, die hier vielversprechend sein könnten, reichen von der Minimierung von Abhängigkeiten in der Software über die Bildung von gekapselten Funktionen bis hin zur Entwicklung von optimierten Systemarchitekturen.

Einen Fortschritt im Bereich Effizienzoptimierung der Ausschlussbeweissführung könnte durch eine Automatisierung der Methode in einem Tool erbracht werden. Dazu ist ein Werkzeug zu entwickeln, das die beschriebenen Ausschlussbeweise softwaretechnisch umsetzt, so dass der manuelle Testauswahlaufwand verringert werden kann. Dies bietet sich insbesondere für den zweiten Ausschlussbeweis (AB II) für Hardware- und Softwareänderungen an, der eine Testzieloptimierung durchführt. Hier eignet sich der Einsatz von Test Suite Reduction Methods (wie z.B. der H-Algorithmus), die aus einer Testmenge eine minimale Untermenge bilden, die bestimmte Anforderungen erfüllt. Die Werkzeugunterstützung ermöglicht dabei zudem die Integration einer Multikriterienoptimierung, wodurch es möglich wird, zusätzlich zur Minimierung der Gesamtanzahl an Tests noch weitere Optimierungskriterien zu berücksichtigen (z.B. Kosten, Kundenvwertigkeit, Ausführungszeit oder Sicherheitsrelevanz der Tests).

Abschließend stellt sich die Frage der Erweiterbarkeit des methodischen Ansatzes, der sich durch eine hohe Sicherheit bei gleichzeitiger Effizienz auszeichnet. Dabei ist zu prüfen, ob das Konzept auf weitere Testebenen oder auf Varianten- bzw. Produktlinien übertragbar ist. Hier ist vor allem der zweite Punkt von Bedeutung, da die ISO 26262 explizit eine Überprüfung der Auswirkungen von Änderungen auf alle betroffenen Varianten fordert (siehe Abschnitt 3.3).

Literaturverzeichnis

Aigner, J.: Zur zuverlässigen Beurteilung von Fahrzeugen. Automobiltechnische Zeitschrift (ATZ), 84(9), 1982, S. 447-450

Albers, A.: Vorwort zu Band 47 der Forschungsberichte des Instituts für Produktentwicklung. Karlsruhe, 2010

Alt, O.: Car Multimedia Systeme Modell-basiert testen mit SysML. Vieweg+Teubner Verlag, Wiesbaden, 2009

Altshuller, G.: Creativity as an exact science. Gordon & Breach, Luxemburg, 1984

AUTOSAR: <http://www.autosar.org>, abgerufen am 18.6.2015

Baller, H.; Lity S.; Lochau, M.; Schaefer, I.: Multi-objective test suite optimization for incremental product family testing. IEEE International Conference on Software Testing, Verification and Validation, 2014

Benda, H. v.: Die Häufigkeit von Verkehrssituationen. FP 7320 im Auftrag der Bundesanstalt für Straßenwesen, Technische Universität München, Lehrstuhl für Psychologie, 1977

Borgeest, K.: Elektronik in der Fahrzeugtechnik. Vieweg+Teubner Verlag, Wiesbaden, 2008

Borgeest, K.: Elektronik in der Fahrzeugtechnik. Vieweg+Teubner Verlag, Wiesbaden, 2010

Broy, M.: Challenges in automotive software engineering. In: Proceedings of the International Conference on Software Engineering (ICSE), Shanghai, 2006

Broy, M.; Reichart, G.; Rothhardt, L.: Architekturen softwarebasierter Funktionen im Fahrzeug: von den Anforderungen zur Umsetzung. In: Informatik Spektrum, Ausgabe 34(1), 2011, S. 42-59

Buchgeher, G.; Ernstbrunner, C.; Ramler, R.; Lusser, M.: Towards tool-support for test case selection in manual regression testing. IEEE International Conference on Software Testing, Verification and Validation, Luxemburg, 2013

Burgdorf, F.: Eine kunden- und lebenszyklusorientierte Produktfamilienabsicherung in der Automobilindustrie. Dissertation, Karlsruher Institut für Technologie, KIT Scientific Publishing, Karlsruhe, 2010

- Caliebe, P.; Herpel, T.; German, R.:** Dependency-based test case selection and prioritization in embedded systems. 5th International Conference on Software Testing, Verification and Validation (ICST), Montreal, 2012
- Chittimalli, P.K.; Harrold, M.J.:** Regression test selection on system requirements. India Software Engineering Conference, 2008
- Clarkson, P.J.; Simons, C.; Eckert, C.:** Predicting change propagation in complex design. In: Proceedings of ASME Design Engineering Technical Conferences and Computers and Information in Engineering Conference, DETC, Pittsburgh, 2001
- Cohen, T.; Navathe, S.B.; Fulton, R.E.:** C-FAR, change favorable representation. Computer-Aided Design, Vol. 32, 2000, S. 321-328
- Deutsches Institut für Normung (DIN):** DIN 69901-1: Projektmanagement – Projektmanagementsysteme – Teil 1: Grundlagen. 2009
- Dickmann, P.:** Schlanker Materialfluss. 2. Auflage, Springer Verlag, Berlin Heidelberg, 2009
- Domsch, C.; Negele, H.:** Einsatz von Referenzfahrsituationen bei der Entwicklung von Fahrerassistenzsystemen. In: 3. Tagung Aktive Sicherheit durch Fahrerassistenz, Garching, 2008
- Drummond, C.; Holte, R.C.:** Explicitly representing expected cost: an alternative to ROC representation. In: Proceedings of the 6th International Conference on Knowledge Discovery and Data Mining, 2000, S. 155-164
- DSM Web:** Design Structure Matrix. <http://www.dsmweb.org/en/understand-dsm/tutorials-overview/descripton-design-structre.html>, abgerufen am 18.06.2015
- Düser, T.:** X-in-the-Loop – ein durchgängiges Validierungsframework für die Fahrzeugentwicklung am Beispiel von Antriebsstrangfunktionen und Fahrerassistenzsystemen. Karlsruher Institut für Technologie, Forschungsberichte des IPEK – Institut für Produktentwicklung, Karlsruhe, 2010
- Eckert, C.; Clarkson, P.J.; Zanker, W.:** Change and customization in complex engineering domains. Research in Engineering Design, Vol. 15, 2004, S. 1-2
- Engström, E.; Skoglund, M.; Runeson, P.:** Empirical evaluations of regression test selection techniques: a systematic review. 2nd International Symposium on Empirical Software Engineering and Measurement (ESEM), Kaiserslautern, 2008
- Engström, E.; Runeson, P.; Skoglund, M.:** A systematic review on regression test selection techniques. Information and Software Technology, 52(1), 2010, S. 14-30
- Engström, E.; Runeson, P.:** A qualitative survey of regression testing practices. In: Ali Babar, M. et al. (Hrsg.): PROFES 2010, LNCS 6265, Springer Verlag, Berlin Heidelberg, 2010, S. 3-16

Engström, E.; Runeson, P.; Ljung, A.: Improving regression testing transparency and efficiency with history-based prioritization – an industrial case study. IEEE Conference on Software Testing, Verification and Validation, 2011

Eppinger, S.D.; Whitney, D.E.; Smith R.P.; Gebala D.A.: A model-based method for organizing tasks in product development. Research in Engineering Design, 1994, S. 1-13

European Commission: ECE R13 H: Uniform provisions concerning the approval of passenger cars with regard to braking. 2014

Eversheim, W.; Bochtler, W.; Laufenberg, L.: Simultaneous Engineering. Springer Verlag, Berlin Heidelberg, 1995

Fastenmeier, W.: Autofahrer und Verkehrssituation – Neue Wege zur Bewertung von Sicherheit und Zuverlässigkeit moderner Straßenverkehrssysteme. Verlag TÜV Rheinland, Köln, Deutscher Psychologen Verlag, Bonn, 1995

Feldhusen, J.; Grote, K.-H. (Hrsg.): Pahl/Beitz Konstruktionslehre – Methoden und Anwendung erfolgreicher Produktentwicklung. 8. Auflage, Springer Vieweg Verlag, Berlin Heidelberg, 2013

Flanagan, T.L.; Eckert, C.; Smith, J.; Eger, T.; Clarkson, P.J.: A functional analysis of change propagation. In: Proceedings of the 14th International Conference on Engineering Design (ICED'03), Stockholm, 2003

Franz, K.: Handbuch zum Testen von Web- und Mobile-Apps. Springer-Verlag, Berlin Heidelberg, 2015

Fricke, E.: Der Änderungsprozess als Grundlage einer nutzerzentrierten Systementwicklung. Dissertation, TU München, 1999

Gallagher, K.; Hall, T.; Black, S.: Reducing regression test size by exclusion. International Conference on Software Maintenance, IEEE, 2007, S. 154-163

Gorthi, R.P.; Pasala, A.; Chanduka, K.K.P.; Leong, B.: Specification-based approach to select regression test suite to validate changed software. 15th Asia-Pacific Software Engineering Conference, 2008

Grande, M.: 100 Minuten für Konfigurationsmanagement. Vieweg+Teubner Verlag, Wiesbaden, 2013

Gustavsson, H.: Architecting automotive product lines. In: Bosch, J., Lee, J. (Hrsg.): SPLC 2010, LNCS 6287, Springer Verlag, Berlin Heidelberg, 2010, S. 92-105

Harrold, M.J.; Gupta, R.; Soffa, M.L.: A methodology for controlling the size of a test suite. ACM Transactions on Software Engineering and Methodology, Ausgabe 2(3), 1993, S. 270-285

- Hirtz, J.; Stone, R.B.; McAdams, D.A.; Szykman, S.; Wood, K.L.:** A functional basis for engineering design: reconciling and evolving previous efforts. *Research in Engineering Design*, Ausgabe 13, 2002, S. 65-82
- Holzmann, H.; Hahn, K.M.; Webb, J.; Mies, O.:** Simulationsbasierte ESP-Homologation für Pkw. *Automobiltechnische Zeitschrift (ATZ)*, Ausgabe 09/2012, 2012, S. 698 – 702
- International Organization for Standardization:** ISO/TS 16949: Quality management systems. 2009
- International Organization for Standardization:** ISO 26262: Road vehicles – Functional Safety. 2011
- International Organization for Standardization:** ISO/IEC 15504: Information technology – Process assessment. 2012
- International Software Testing Qualifications Board (ISTQB):** Glossary. <http://astqb.org/glossary/search/actual%20result>, abgerufen am 02.07.2015
- IPG:** Car Maker. <http://ipg.de/simulationsolutions/carmaker/>, abgerufen am 22.6.2015
- Jarratt, T.; Eckert, C.M.; Clarkson, P.J.:** Development of a product model to support engineering change management. In: *Proceedings of the TCME 2004*, Lausanne, 2004
- Juergens, E.; Hummel, B.; Deissenboeck, F.; Feilkas, M.; Schlögel, C.; Wübbeke, A.:** Regression test selection of manual system tests in practice. 15th European Conference on Software Maintenance and Reengineering. Oldenburg, 2011
- Kaner, C.:** Improving the maintainability of automated test suites. In: *Proceedings of Quality Week*, 1997
- Keller, R.; Eger, T.; Eckert, C.M.; Clarkson, P.J.:** Visualizing change propagation. *International Conference on Engineering Design (ICED'05)*, Melbourne, 2005
- Khan, S.; Rehman, O.; Malik, S.:** The impact of test case reduction and prioritization on software testing effectiveness. In: *International Conference on Emerging Technologies, ICET*, 2009, S. 416-421
- Kichigan, D.:** A method for test suite reduction for regression testing of interactions between software modules. In: Pnueli, A. et al. (Hrsg.): *PSI 2009, LNCS 5947*, Springer Verlag, Berlin Heidelberg, 2010, S. 177-184
- Kim, J.; Porter, A.; Rothermel, G.:** An Empirical study of regression test application frequency. In: *Software Testing, Verification and Reliability*, Vol. 15, 2005, S. 257-279
- Klem, A.; Schlicht, O.; Brand, M.; Gerlinger, U.:** Erprobung der Langzeitqualität des neuen Audi A4. In: *Sonderausgabe von ATZ und MTZ – ATZ extra Audi A4*, September 2007, S. 62-67

Kohn, W.: Statistik – Datenanalyse und Wahrscheinlichkeitsrechnung. Springer Verlag, Berlin Heidelberg, 2005

Kvasnicka, P.; Prokop, G.; Dörle, M.; Rettinger, A.; Stahl, H.: Durchgängige Simulationsumgebung zur Entwicklung und Absicherung von fahrdynamischen Regelsystemen. In: Berechnung und Simulation im Fahrzeugbau, Würzburg, VDI-Berichte 1976, VDI-Verlag, Düsseldorf, 2006, S. 387-404

Langermann, R.: Beitrag zur durchgängigen Simulationsunterstützung im Entwicklungsprozess von Flugzeugsystemen. Dissertation, TU Braunschweig, 2008

Lehnert, S.: A review of software change impact analysis. TU Ilmenau, Report ilm1-2011200618, 2011

Leung, H.; White, L.: A cost model to compare regression test strategies. In: Proceedings of the Conference on Software Maintenance, 1991, S. 201-208

Leung, H.; White, L.: A study of integration testing and software regression at the integration level. In: Microsoft Research (Hrsg.) – TechReport MSR-TR-2005-94, 2005

Liggismeyer, P.: Software-Qualität – Testen, Analysieren und Verifizieren von Software. 2. Auflage, Spektrum Akademischer Verlag, Heidelberg, 2009

Lindemann, U.; Reichwald, R.: Integriertes Änderungsmanagement. Springer Verlag, Berlin Heidelberg, 1998

Malishevsky, A.G.; Rothermel, G.; Elbaum, S.: Modeling the cost-benefits tradeoffs for regression testing techniques. In: Proceedings of the International Conference on Software Maintenance, 2002

Mao, Y.; Wiessalla, J.; Meier, J.; Risse, W.; Mathot, G.; Blum, M.: CAE supported ESC development/release process. In: Proceedings of the FISITA 2012 World Automotive Congress, 2012

Mechanical Simulation: CarSim. <http://carsim.com/products/carsim/index.php>, abgerufen am 22.06.2015

Meywerk, M.: CAE-Methoden in der Fahrzeugtechnik. Springer Verlag, Berlin Heidelberg, 2007

Milling, P.; Jürging, J.: Der Serienanlauf in der Automobilindustrie. In: Himpel, F. et al. (Hrsg.): Spektrum des Produktions- und Innovationsmanagements. Gabler Verlag, Wiesbaden, 2008, S. 67-82

Mokammel, F.: Impact analysis of graph-based requirements models using PageRank algorithm. In: Proceedings of the 2013 IEEE International Systems Conference, 15.-18. April, Orlando, 2013

- National Highway Traffic Safety Administration (NHTSA):** Federal Motor Vehicle Safety Standard (FMVSS) No. 126: Electronic Stability Control Systems. 2007
- Neale, V.L.; Klauer, S.G.; Knipling, R.R.; Dingus, T.A.; Holbrook, G.T.; Petersen, A.:** The 100 car naturalistic driving study, Phase 1 – Experimental design. NHTSA Report DOT HS809 536, 2002
- Neukum, A.; Krüger, H.:** Fahrerreaktionen bei Lenksystemstörungen – Untersuchungsmethodik und Bewertungskriterien. In: Reifen, Fahrwerk, Fahrbahn. VDI-Berichte 1791, VDI-Verlag, Düsseldorf, 2003
- Nörenberg, R.:** Effizienter Regressionstest von E/E-Systemen nach ISO 26262. KIT Scientific Publishing, Karlsruhe, 2012
- Object Mangament Group (OMG):** Unified Modeling Language (UML). <http://www.uml.org/>, abgerufen am 18.6.2015
- Orso, A.; Do, H.; Rothermel, G.; Harrold, M. J.; Rosenblum, D.S.:** Using component metadata to regression test component-based software. Software Testing, Verification and Reliability, Vol. 17, No. 2, 2007, S. 61-94
- Parsa, S.; Khalilian, A.; Fazlalizadeh, Y.:** A new algorithm to test suite reduction based on cluster analysis. 2nd International Conference on Computer Science and Information Technology (ICCSIT), Beijing, 2009
- Paulic, A.; Weiss, G.; Eilers, D.:** Verbessertes Softwaretest für Steuergeräte kombiniert sequenzbasierten und modellbasierten Test. ATZ elektronik, Ausgabe 2/2012, 2012, S. 150-154
- Popper, K.R.:** Logik der Forschung. 10. Auflage, Mohr Siebeck, Tübingen, 2005
- Raffaelli, R.; Germani, M.; Graziosi, S.; Mandorli, F.:** Development of a multilayer change propagation tool for modular products. International conference on engineering design, ICED, Paris, 2007
- Reichart, G.:** Menschliche Zuverlässigkeit beim Führen von Kraftfahrzeugen. VDI-Verlag, Düsseldorf, 2001
- Reif, K.:** Automobilelektronik. ATZ/MTZ-Fachbuch, Vieweg Verlag, Wiesbaden, 2009
- Reuter, A.:** Produkthaftung in Deutschland. In: Werdich, M. (Hrsg.): FMEA – Einführung und Moderation. Vieweg+Teubner Verlag, Wiesbaden, 2011
- Ring, J.; Fricke, E.:** Rapid evolution of all your systems – problem or opportunity. 17th DASC, IEEE, 1998
- Roddeck, W.:** Einführung in die Mechatronik. Vieweg+Teubner, Wiesbaden, 2012
- Rothermel, G.:** Efficient, effective regression testing using safe test selection techniques. Dissertation, Clemson University, 1996

Rothermel, G.; Harrold, M. J.: Analyzing regression test selection techniques. IEEE Transactions on Software Engineering, Vol. 22, No. 8, 1996, S. 529-551

Rothermel, G.; Harrold, M.: Empirical studies of a safe regression test selection technique. IEEE Transactions of Software Engineering, 1998

Rothermel, G.; Harrold, M.J.; von Ronne, J.; Hong, C.: Empirical studies of test-suite reduction. In: Software Testing, Verification and Reliability. Vol. 12, No. 4, 2002, S. 219-249

Runkler, T. A.: Data Mining – Methoden und Algorithmen intelligenter Datenanalyse. Vieweg+Teubner Verlag, Wiesbaden, 2010

Sajeev, A.; Wibowo, B.: Regression test selection based on version changes of components. In: Tenth Asia-Pacific Software Engineering Conference, 2003, S. 78-85

Schäuffele, J.; Zurawka, T.: Automotive Software Engineering, ATZ/MTZ Fachbuch, Wiesbaden, 2003

Schramm, D.; Hiller, M.; Bardini, R.: Modellbildung und Simulation der Dynamik von Kraftfahrzeugen. 2. Auflage, Springer Vieweg Verlag, Berlin Heidelberg, 2013

Schuh, G.; Müller, J.; Rauhut, M.: Gestaltung von Produktentwicklungsprozessen. In: Schuh, G. et al. (Hrsg.): Innovationsmanagement 3. Springer-Verlag, Berlin Heidelberg, 2012

Schulz, A.; Fricke, E.: Incorporating flexibility, agility, robustness, and adaptability within the design of integrated systems. In: Proceedings of the 18th Digital Avionics Systems Conference, 1999

Schulz, A.; Fricke, E.; Igenbergs, E.: Enabling changes in systems throughout the entire life-cycle. In: Proceedings of the 10th annual INCOSE Conference, INCOSE, Minneapolis, 2000

Siegl, S.; Hielscher, K.S.; German, R.; Berger, C.: Formal specification and systematic model-driven testing of embedded automotive systems. Design, Automation and Test in Europe Conference (DATE), Grenoble, 2011

Starke, G.: Effektive Softwarearchitekturen – ein praktischer Leitfaden. 5. Auflage, Hanser Verlag, München, 2011

Steward, D.: The design structure system: a method for managing the design of complex systems. IEEE Transactions on Engineering Management, Vol. EM-28, No. 3, 1981, S. 71-74

Stone, R.; Wood, K.: Development of a functional basis for design. Proceedings of DETC99, Las Vegas, 1999

- Struss, P.; Fraracci, A.:** Automated model-based FMEA of a braking system. 8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (Safeprocess), Mexico City, 2012
- Sundmark, D.; Petersen, K.; Larsson, S.:** An exploratory case study of testing in an automotive electrical system release process. In: Proceedings of the 6th IEEE International Symposium on Industrial Embedded Systems (SIES), Västerås, 2011
- Szykman, S.; Racz, J.W.; Sriram, R.D.:** The representation of function in computer-based design. Proceedings of the 1999 ASME Design Engineering Technical Conferences, Las Vegas, 1999
- Tan, P.-N.:** Receiver operating characteristic. In: Liu, L.; Özsu, M.T. (Hrsg.): Encyclopedia of database systems. Springer, 2009, S. 2349-2352
- TESIS:** veDYNA. <http://www.thesis-dynaware.com/produkte/vedyna/uebersicht.html>, abgerufen am 22.6.2015
- Ungermann, J.:** Zuverlässigkeitsnachweis und Zuverlässigkeitsentwicklung in der Gesamtfahrzeugerprobung. Dissertation, ETH Zürich, 2009
- Verband der Automobilindustrie (VDA):** Automotive SPICE Prozessassessment. 1. Auflage, 2007
- Vokolos, F.; Frankl, P.:** Pythia: A regression test selection tool based on text differencing. In: Proceedings of the International Conference on Reliability, Quality, and Safety of Software Intensive Systems, 1997
- Wannenwetsch, H.:** Integrierte Materialwirtschaft und Logistik. 2. Auflage, Springer Verlag, Berlin Heidelberg, 2004
- Weitzel, A.:** Objektive Bewertung der Kontrollierbarkeit nicht situationsgerechter Reaktionen umfeldsensorbasierter Fahrerassistenzsysteme. VDI Fortschritts Berichte, Reihe 12, Nr. 774, VDI Verlag, Düsseldorf, 2013
- Westermann, T.:** Modellbildung und Simulation. Springer Verlag, Berlin Heidelberg, 2010
- Westkämper, E.:** Null-Fehler-Produktion in Prozessketten. Springer Verlag, Berlin Heidelberg, 1997
- Wilhelm, U.; Ebel, S.; Weitzel, A.:** Funktionale Sicherheit und ISO 26262. In: Winner H. et al. (Hrsg.): Handbuch Fahrerassistenzsysteme, 3. Auflage, Springer Vieweg Verlag, Berlin Heidelberg, 2015, S. 87-105
- Yazdanshenas, A.R.; Moonen, L.:** Fine-grained change impact analysis for component-based product families. In: International Conference on Software Maintenance (ICSM), IEEE, 2012

Yoo, S.; Harman, M.: Regression testing minimization, selection and prioritization: a survey. *Software Testing, Verification and Reliability*, 2010

Zanner, S., Jäger, S., Stotko, C.M.: Änderungsmanagement bei verteilten Standorten. In: *Industrie Management*, Ausgabe 18/3, 2002, S. 40-43

Zhao, O.; Yan, H.; Xiang, L.; Xu, B.: Change impact analysis to support architectural evolution. *Journal of Software Maintenance: Research and Practice – Special Issue: Separation of Concerns for Software Evolution*, 2002

Zheng, J.; Robinson, B.; William, L.; Smiley, K.: An initial study of a lightweight process for change identification and regression test selection when source code is not available. *16th International Symposium on Software Reliability Engineering*, 2005

Eigene Veröffentlichungen

Singer, C.; Vey, C.; Winner, H.: Systematik der Gesamtfahrzeugabsicherung und Verbesserungspotenziale. In: Zeitschrift für die gesamte Wertschöpfungskette Automobilwirtschaft (ZfAW), Ausgabe 1/2015, FAW-Verlag, Bamberg, 2015

Winner, H.; Hakuli, S.; Lotz, F.; Singer C. (Hrsg.): Handbuch Fahrerassistenzsysteme. ATZ/MTZ-Fachbuch, Springer Fachmedien, Wiesbaden, 2015

Winner, H.; Hakuli, S.; Lotz, F.; Singer C. (Hrsg.): Handbook of Driver Assistance Systems. Springer International Publishing, Switzerland, 2016

Betreute studentische Arbeiten

Adam, D.: Messdatenbasierte Parametrisierung eines Fahrzeugmodells zur Fahrdynamiksimulation mit IPG CarMaker. Bachelor Thesis 1129/12, 2012

Arnemann, P.: Entwicklung eines Konzeptes zur Beobachtbarkeit von Fehlapplikationen der Funktion Kurvenbremsung eines Bremsregelsteuergeräts. Master Thesis Nr. 515/12, 2013

Schädlich, C.: Analyse von Indikatoren zur Beschreibung der Freigabequalität von Bremsregelsteuergeräten. Studienarbeit Nr. 1183/14, 2014

Schmidt, L.: Analyse der Bewertungskriterien bei der Applikation einer Bremsregel-funktion. Bachelor Thesis 1149/13, 2013

Schöler, N.: Analyse von Indikatoren zur Beschreibung des Applikations- und Freigabeaufwandes von Bremsregelsteuergeräten. Master Thesis 530/13, 2014

Seibel, M.: Entwicklung und Implementierung eines Konzeptes zur Identifikation der Schwerpunkthöhe eines Fahrzeugs. Master Thesis 505/12, 2012

Weidenfeller, C.: Entwicklung von Kriterien für die Identifikation von Applikationsparametern der Funktion „Kurvenbremsen“ des Bremsregelsteuergeräts. Bachelor Thesis 1148/13, 2013

Lebenslauf

Persönliche Daten

Name: Christina Theresia Singer
Geburtsdatum: 15.10.1983
Geburtsort: Lippstadt

Schulzeit

1994 – 2003 Friedrich-Spee-Gymnasium Rüthen

Studium

2005 - 2008 Studium Allgemeiner Maschinenbau
Fachhochschule Südwestfalen Meschede

2008 - 2011 Studium Allgemeiner Maschinenbau
Technische Universität Darmstadt

Berufstätigkeit

2011 – 2015 Wissenschaftliche Mitarbeiterin am Fachgebiet Fahrzeugtechnik der Technischen Universität Darmstadt