

Praxis der Forschung – Sommersemester 2016

Teilnehmende Lehrstühle im SoSe 2016

- ITI Prof. Beckert, Arbeitsgruppe für Anwendungsorientierte formale Verifikation
- TM Prof. Beigl, Lehrstuhl für Pervasive Computing Systems / TECO
- IPD Prof. Reussner / Jun.-Prof. Koziolk, Software Design and Quality (SDQ)
- IAR Prof. Asfour, Arbeitsgruppe für Hochperformante Humanoide Technologien (H²T)

Kontakt bei allgemeinen Fragen zu „Praxis der Forschung“:

- Sarah Grebing, ITI Prof. Beckert, sarah.grebing@kit.edu, +49 721 608-45253
- Matthias Budde, TM Prof. Beigl, matthias.budde@kit.edu, +49 721 608-41705
- Anja Bachmann, TM Prof. Beigl, bachmann@teco.edu, +49 721 608-41713

Termine:

- Anmeldung bis 26.04.16 beim jeweiligen Betreuer, per Mail bei sarah.grebing@kit.edu und im ILIAS Kurs [Praxis der Forschung \(Methoden, 1. Semester\) SoSe 16](https://ilias.studium.kit.edu/goto_produkativ_crs_527444.html)
https://ilias.studium.kit.edu/goto_produkativ_crs_527444.html
Bitte Thema und Betreuer bei der Anmeldung angeben.
- 1. Methodische Veranstaltung:
„Kickoff für das SoSe 2016 und Literaturrecherche“
29.04.2016, 14:00 - 15:30 Uhr in Raum 010, Geb. 50.34

Ausgeschriebene Themen im SoSe 2016

Testen von Informationsflusseigenschaften.....	2
Deep Learning for Semantic Scene and Action Perception.....	2
Scalable Anomaly Detection for Smart City Infrastructure Networks	3
Real-Time Identification and Characterization of Events in Social Media	3
Interruptibility Detection on Mobile Devices	4
Nichtinterferenz in relationalen Datenbanksystemen	4
Qualitätsmodellierung von Subsystemen bei architekturellen Entwurfsentscheidungen	5
Inkrementelle Modellsichten	5

Testen von Informationsflusseigenschaften

Es ist wünschenswert, dass Programme sensible Informationen wie Kontodaten oder Passwörter nicht veröffentlichen. Um dies sicherzustellen, kann man mit Hilfe von Verifikationswerkzeugen die Datenflüsse eines Programms analysieren und zeigen, dass keine sicherheitsrelevante Informationen geleckt werden. Die Eingaben und Ausgaben des Programms werden in zwei Kategorien eingeteilt – vertraulich (high) und öffentlich (low). Man muss dann zeigen, dass die öffentlichen Ausgaben unabhängig von den vertraulichen Eingaben sind, also bei gleichen öffentlichen Eingaben müssen gleiche öffentliche Ausgaben herauskommen, unabhängig von den vertraulichen Eingaben. Diese Eigenschaft (non-interference) spricht also über zwei beliebige Programmabläufe und ist deswegen schwieriger zu beweisen als funktionale Eigenschaften, die nur einen Programmablauf betrachten. Da das Beweisen dieser Eigenschaften oft zu kompliziert ist, möchten wir sie testen. Dafür müssen ausführbare Testfälle aus einem Teilbeweis generiert werden. Das Testen erfolgt automatisch und ist somit einfacher als das Beweisen, wo interaktive Schritte nötig sein können. Allerdings kann ein Test die gewünschte Eigenschaft nur für die beim Testen verwendete Eingaben zeigen. Daher müssen die Testeingaben so ausgewählt werden, dass sie das Programmverhalten so weit wie möglich abdecken.

Mögliche Aufgaben im Rahmen dieses Themas sind:

- 1) Definieren des Konzepts eines Informationsfluss-Testfalls
- 2) Definieren von Abdeckungskriterien für Informationsfluss-Testfälle, wie aussagekräftig ist eine Testsuite?
- 3) Erweiterung des Testfallgenerierung-Frameworks in KeY, so dass neben funktionalen jetzt auch Informationsflusseigenschaften getestet werden können
- 4) Exploit-Generation - statt eine Testsuite zu generieren, die möglichst viel abdeckt, gezielt nach Inputs suchen, die die Eigenschaft verletzen und Testfälle nur für diese Inputs erzeugen. Die generierte Testsuite soll einen Angreifer simulieren, der die öffentliche Eingabe beliebig setzen kann, um Informationen über die vertrauliche Eingabe zu gewinnen.

Kontakt / Betreuer: Mihai Herda (ITI Beckert) herda@kit.edu

Deep Learning for Semantic Scene and Action Perception

When a robot needs to imitate an observed action sequence, it must first understand the inherent characteristic features of individual actions that remain the same even under different circumstances. These descriptive features reflect the semantics of actions which has high degree of invariance between different demonstrations of the same action in various scene contexts. In this sense, the more descriptive and structural the action semantics, the greater the capabilities and autonomy of the robot. In this study, we aim at applying deep learning techniques to allow our humanoid robot ARMAR-IIIa to explore the semantics of the scene and actions by processing the perceived raw RGB-D image streams. The proposed work will yield localization and recognition of possible functional areas from an unseen environment.

In this project, the candidate has to conduct various experiments on already existing publicly available datasets. First, a multi-layer Convolutional Neural Network (CNN) has to be implemented for scene functionality analysis. To train the network, different existing labeled datasets will be considered. The final performance of the system has to be tested with our humanoid robot ARMAR-IIIa in a kitchen environment. Applicants must be interested in deep learning and should possess good programming skills either in Python, C/C++, or Matlab.

Kontakt / Betreuer: Eren Erdal Aksoy (IAR Asfour) eren.aksoy@kit.edu

Scalable Anomaly Detection for Smart City Infrastructure Networks

Internet-enabled, location aware smart phones with sensor inputs have led to novel urban monitoring applications exploiting unprecedented high levels of citizen participation in dense metropolitan areas [2]. For policy makers, it is a key task to keep track of trends and developments for understanding and effectively reacting to problems around a city, specially in their early stages. Anomaly and Event Detection Approaches (see [1]) help understand the urban dynamics contained in this kind of data and detect anomalies and new developments of urban dynamics, a key-task in today s Smart Cities.

Goal of this work is to evaluate the application and scalability of anomaly detection algorithms to (big) urban city data. This includes the prototypical implementation of state of the art algorithms (see [1]) and its evaluation in a smarter city context on a large number of data sources.

[1] Chen, F., & Neill, D. B. (2014, August). Non-parametric scan statistics for event detection and forecasting in heterogeneous social media graphs. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1166-1175). ACM.

[2] Budde, M., De Melo Borges, J., Tomov, S., Riedel, T., & Beigl, M. (2014, October). Leveraging spatio-temporal clustering for participatory urban infrastructure monitoring. In Proceedings of the First International Conference on IoT in Urban Space (pp. 32-37). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Kontakt / Betreuer: Julio Borges (TM Beigl) borges@teco.edu

Real-Time Identification and Characterization of Events in Social Media

With the increasing volume of location-annotated content from various social media platforms like Twitter, Instagram and Foursquare, we now have real-time access to people s daily documentation of local activities, interests and attention.

At the same time, much of the content on Social Media does not correspond to any particular trending event, making the separation between trending event and non-event content challenging and essential for the unknown event identification task.

Event detection from social media is a vibrant research area that draws on techniques from various fields such as machine learning, natural language processing, data mining, information extraction and retrieval, and text mining.

Goal of this project is to characterize and classify the different techniques which can be applied for real-time event detection providing an exhaustive review of existing approaches (broad survey). Finally, a few techniques should be selected and benchmarked to evaluate the performance of different detection approaches and various features on both synthetical and real datasets.

[1] Xia, C., Schwartz, R., Xie, K. E., Krebs, A., Langdon, A., Ting, J., & Naaman, M. (2014, April). Citybeat: Real-time social media visualization of hyper-local city data. In Proceedings of the companion publication of the 23rd international conference on World wide web companion (pp. 167-170). International World Wide Web Conferences Steering Committee.

[2] Atefeh, F., & Khreich, W. (2015). A survey of techniques for event detection in twitter. Computational Intelligence, 31(1), 132-164.

Kontakt / Betreuer: Julio Borges (TM Beigl) borges@teco.edu

Interruptibility Detection on Mobile Devices

Vor allem im Bereich des Experience Samplings, der Ecological Momentary Analysis oder des ambulanten Assessments sind Abfragen an Nutzer teilweise unvermeidlich, aber sehr wichtig. Wo früher noch über Erinerungen und Wecker an das Ausfüllen von analogen Fragebögen oder das Schreiben von Tagebüchern erinnert wurde, so erfolgt heutzutage alles digital.

Über Notifications werden die Nutzer gebeten, im Rahmen einer App ihre Angaben vorzunehmen. Problematisch ist es nur, wenn man in einer ungünstigen Situation die Notification schickt, sie zu penetrant ist (z.B. lautes Klingeln oder Vibrieren) oder zu häufig Notifications gesendet werden. Dies kann dazu führen, dass Nutzer aus Studien aussteigen oder in der Schnelle der Situation falsche Angaben machen. Dies soll natürlich vermieden werden.

Fraglich ist jedoch einerseits, wann man den Nutzer am Mobilgerät am ehesten stören kann und andererseits wie lange kann man eine Notification problemlos hinauszögern, sodass sich der Nutzer noch erinnert. Wir wollen diese „Interruptibility“ in verschiedenen Kontexten untersuchen!

Aufgaben:

- Führen von Interviews oder Umfragen zum Erfassen passender Kontexte
- Implementieren von Klassifikatoren für diese Kontexte („Interruptability Detectors“) unter Android unter Nutzung verfügbarer Sensoren und APIs
- Evaluation im Rahmen einer Nutzerstudie

Kontakt / Betreuer: Anja Bachmann (TM Beigl) bachmann@teco.edu

Nichtinterferenz in relationalen Datenbanksystemen

Die Geheimhaltung von Informationen in modernen IT Systemen gewinnt auf Grund von immer wachsender Mengen an Daten stets an Wichtigkeit. Um diese Eigenschaft in Web-basierten Anwendungen garantieren zu können haben sich Methoden etabliert, die den Fluss von Informationen durch Programme analysieren. Typischerweise wird hierbei untersucht, inwiefern geheimzuhaltende Eingaben in ein Programm öffentliche Ausgaben beeinflussen. Beeinflussen geheime Eingaben die öffentlichen Ausgaben gar nicht, so ist das Programm nichtinterferent.

In den letzten Jahren wurden große Fortschritte in der Spezifikation und Analyse von verteilten Systemen bezüglich solcher Nichtinterferenzeigenschaften gemacht. Allerdings wurden dabei zentrale Technologien von Webanwendungen, insbesondere relationale Datenbanken, bisher nicht betrachtet. Das Verhalten von relationalen Datenbanken ist typischerweise definiert durch SQL-Anfragen, die wiederum durch relationale Algebra formal beschrieben werden können. Dadurch bieten relationale Datenbanken alle Voraussetzungen, die notwendig sind, um Nichtinterferenz auch in diesen Systemen zu analysieren und formal nachzuweisen.

Ziel dieser Projektgruppe ist es bestehende Nichtinterferenzbegriffe für verteilte Programme auf relationale Datenbanksysteme anzuwenden. Hierfür soll der Nichtinterferenzbegriff geeignet formalisiert werden und eine geeignete Methode zur Spezifikation von Geheimnissen und öffentlichen Informationen in Datenbanken entworfen werden. Abschließend soll aufbauend auf Beweissysteme (z. B. SMT Beweiser) eine Methode entwickelt werden, die für ein bestehendes relationales Datenbanksystem und den dazugehörigen Schnittstellendefinitionen untersuchen kann, ob der Datenbankentwurf die spezifizierten Sicherheitseigenschaften erfüllt.

Voraussetzungen: Interesse an Formalen Systemen, Logik, relationaler Algebra und theoretischen Arbeiten

Kontakt / Betreuer: Simon Greiner (ITI Beckert) simon.greiner@kit.edu

Qualitätsmodellierung von Subsystemen bei architekturellen Entwurfsentscheidungen

In modernen Systemen werden bei ihrem Entwurf immer häufiger bereits fertige Systeme als Teil der eigenen Gesamtlösung eingesetzt. Dies sieht sich nicht mehr auf etwa Datenbankmanagementsysteme beschränkt, sondern lässt sich heute durch eine Vielfalt von bereits fertigen Lösungen, die als Subsysteme angesehen werden können, wie Authentifizierungssysteme oder Zahlungslösungen, erweitern. Jede dieser Subsysteme stellt zum einen eine Reihe verschiedener Funktionalitäten zur Verfügung, beeinflusst aber gleichzeitig auch die resultierende Qualität (wie Performance, Zuverlässigkeit, Sicherheit, Bedienbarkeit) der Softwarearchitektur. Einflüsse auf die Qualität ergeben sich zum Einen natürlich aus der Lösung selbst, aber zum anderen auch durch Freiheitsgrade darin, wie die Lösung konkret in die Ziel-Softwarearchitektur eingebaut wird, d.h. wie bereits existierende Komponenten mit der neuen Lösung in Beziehung gebracht werden. Das Ziel dieser Arbeit ist die Modellierung des Qualitätseinflusses durch diese Beziehungen der neuen Komponenten mit den übrigen des Zielsystems. Hierfür sollen zunächst architekturelle Freiheitsgrade untersucht werden und anschließend in Beziehung zu ihrem Einfluss auf die Gesamtqualität der Softwarearchitektur gebracht werden.

Kontakt / Betreuer: Axel Busch (PD Prof. Reussner / Jun.-Prof. Koziolok) abusch@ira.uka.de

Inkrementelle Modellsichten

Moderne Software-Systeme weisen eine hohe Komplexität und eine umfangreiche Größe auf. Daher werden in der Entwicklung solcher Systeme mehrere Sprachen und Modelle verwendet, um unterschiedliche Gesichtspunkte und Abstraktionsebenen der Systeme zu beschreiben. Beispielsweise können Softwaresysteme unter anderem mit Komponentenmodellen der Architektur, Performance-Modellen zur Analyse der nichtfunktionalen Eigenschaften oder mit der Implementierung im Code dargestellt werden. Obwohl all diesen Artefakten unterschiedliche Konzepte und Formalismen zugrunde liegen, beschreiben sie dasselbe System aus unterschiedlichen Blickwinkeln. Zwischen diesen Modellen kommt es häufig zu semantischen Überlappungen, beispielsweise sind Komponenten aus der Architektur auch in der Implementierung in Form von Klassen wieder zu finden. Um ein vollständiges Bild einer Komponente zu erfassen, müssen Informationen daher aus verschiedenen Modellen aggregiert werden in sog. Sichten. Die Entwicklung dieser Sichten wird dabei von Sprachen wie beispielsweise dem am Lehrstuhl Reussner entwickelten ModelJoin unterstützt. Ein grundsätzliches Problem hierbei ist es jedoch, dass sich während des Softwareentwicklungsprozesses die zugrundeliegenden Modelle der Architektur, der Performance-Modelle und der Implementierung ändern. Da diese Änderungen potentiell Auswirkungen auf die Sichten haben, müssen diese neu generiert werden. Für große Softwaresysteme ist die ständige Neugenerierung dieser Sichten aber zeit- und rechenintensiv, gerade weil mit jeder Änderung nur kleine Teile des Systems geändert werden. Es ist daher lohnenswert, möglichst große Teile der bestehenden Sicht zu belassen und nur die Teile zu aktualisieren, die tatsächlich von einer Änderung betroffen sind. Man spricht hierbei von einer inkrementellen Ausführung der Sicht.

Im Rahmen der Projektgruppe soll untersucht werden, inwiefern neue Forschungsansätze zu impliziter Inkrementalität dazu geeignet sind, eine solche inkrementelle Ausführung von Sichten automatisch aus der Spezifikation eines Sichttyps abzuleiten. Dazu soll ein System entwickelt werden, das mit der Sprache ModelJoin spezialisierte flexible Sichten automatisch inkrementelle ausführen kann. Hierbei kann auf das ebenfalls am Lehrstuhl Reussner entwickelte inkrementelle Ausführungssystem NMF Expressions zurückgegriffen werden. Die Korrektheit des Systems soll durch Fallstudien belegt und deren Nützlichkeit durch Messung des Laufzeitverhaltens und des Speicherverbrauchs belegt werden. Studierende werden dabei durch Mitarbeiter des Lehrstuhls Reussner betreut.

Kontakt / Betreuer: Erik Burger (SDQ Reussner) burger@kit.edu
Georg Hinkel (SDQ Reussner) georg.hinkel@kit.edu