SOAMED Coherence Report 2011

SOAMED Ph.D. students

May 7, 2012

1 The big picture

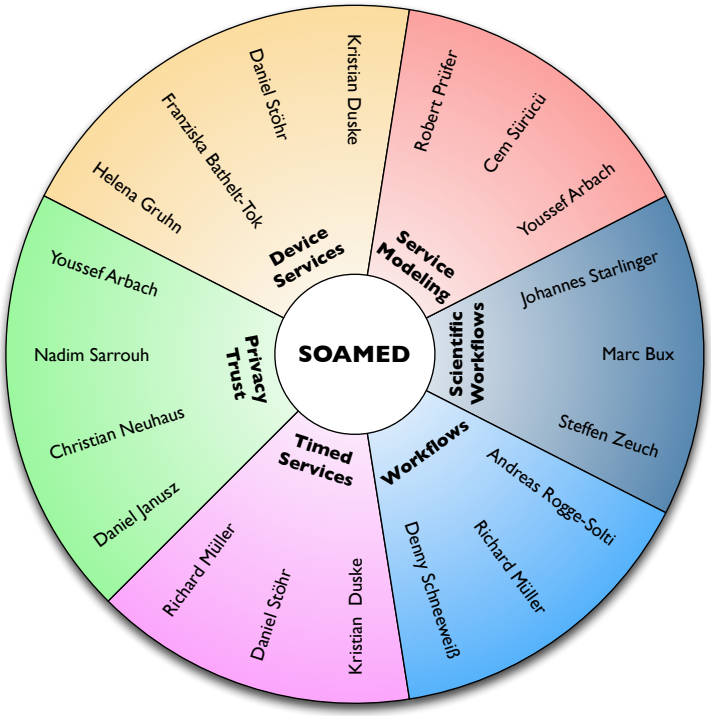

Figure 1: The fields of interest of the SOAMED Ph.D. students.

The graduate school SOAMED centers around the idea to introduce theoretically well-founded service-oriented concepts to healthcare systems and medical devices. This approach aims at a decisive improvement of concepts, methods, and tool support for service-oriented system construction. Currently, sixteen Ph.D. students and eleven advisors participate in SOAMED. For a better overview of the research of the Ph.D. students, we coarsely classified their fields of interest into six parts; that is, device services, service

modeling, scientific workflows, workflows, timed services, as well as privacy and trust. See Fig. 1 for a visualization. Helena Gruhn, Franziska Bathelt-Tok, Daniel Stöhr and Kristian Duske focus on service-oriented concepts for medical devices. The modeling of healthcare services is in the field of interest of Robert Prüfer, Cem Sürücü, and Youssef Arbach. Andreas Rogge-Solti, Richard Müller and Denny Schneeweiss are involved in research about healthcare workflows in general, whereas Johannes Starlinger, Marc Bux and Steffen Zeuch specifically target scientific workflows. Finally, Youssef Arbach, Nadim Sarrouh, Christian Neuhaus and Daniel Janusz investigate privacy and trust in service-oriented systems.

The sixteen Ph.D. students can be split into two generations according to when they started their research in SOAMED. In order to support mentoring of the second generation by the first generation, we have established a weekly SOAMED Ph.D. students meeting at the Humboldt Graduate School since mid 2011. Attendance is voluntary. The Ph.D. students meeting gives all Ph.D. students the opportunity to closely cooperate in related topics, talk about papers or contributions important for the idea of SOAMED, practice presentations for upcoming workshops or conferences, or administrate their teamwork in general.

In the following, each Ph.D. student gives a brief overview of how his research integrates into the big picture of SOAMED, and describes ongoing or planned cooperations with other SOAMED Ph.D. students. A report of Helena Gruhn is missing because she was on maternity leave in 2011.

2 Individual cooperation reports

2.1 Franziska Bathelt-Tok

The research topic of Franziska Bathelt-Tok's dissertation deals with the question, how the interoperability between various medical devices can be enabled. The issue of interoperability is also regarded within the dissertations of Helena Gruhn, Kristian Duske and Daniel Stöhr.

While Kristian Duske and Daniel Stöhr look at the control flow level of device services and try to include time constraints with help of timed petri nets or timed automata,respectively, Franziska Bathelt-Tok and Helena Gruhn focus on the data flows regarding the interactions of device services. In contrast to Helena Gruhn who analyses the characteristics of device services in sensor-actuator-networks, Franziska Bathelt-Tok deals with how to determine message formats of the device services.

2.2 Daniel Stöhr

The topic of Daniel Stöhr's work is the "Automated Composition of Timed Services for Synchronizing Medical Devices". The aspect of timed services is related to the works of Kristian Duske and Richard Müller, while the aspect of device services is related to the works of Franziska Bathelt-Tok, Kristian Duske, and Helena Gruhn.

The link between Daniel Stöhr, Kristian Duske, and Richard Müller is the analysis of formal service models containing timed properties. However, the focus of Daniel Stöhr's work is on the generation of service compositions described as timed automata. In contrast, the other two works deal with how services, represented as timed petri nets, can be analyzed regarding properties like controllability, correctness and conformance.

Moreover, Kristian Duske and Daniel Stöhr perform a joint case study at the Charité Berlin, where they model devices taking part in a diagnostic procedure for congenital hyperinsulinism. Finally, device services are an interlink to the works of Franziska Bathelt-Tok and Helena Gruhn. While Daniel Stöhr investigates the interaction of device services on a control flow level, the former deals with device services on a data flow level. The latter examines the specifics of device services in sensor-actuator-networks.

2.3 Kristian Duske

Kristian Duske's research focuses on the run-time correctness verification. Specifically, the goal of his work is to develop a method to verify functional correctness of a service composition at run-time in the presence of timing information. Since this dissertation is in the area of formal verification, it is related to the works of Daniel Stöhr, Richard Müller and, to a lesser extent, the work of Robert Prüfer. The problem of medical device interoperability provides a good application for run-time correctness verification, so this work is also related to the topics of Franziska Bathelt-Tok and Helena Gruhn.

There is a close link between Daniel Stöhr's work and Kristian Duske's work because both deal with timing properties of services. While both dissertations focus on different aspects of this field, the link is close enough to perform a joint case study at the Charité Berlin, which is already described above.

Kristian Duske has also worked closely with Richard Mülller on the topic of timed services. Together, they have published a paper at the ZEUS Workshop 2012 that surveys approaches on the verification of timed services. Furthermore, since both Richard Müller's and Robert Prüfer's work is based on Petri nets, there are many opportunities to cooperate further in the future.

2.4 Robert Prüfer

The topic of Robert Prüfer's dissertation, "Scenario-based Design of Data-dependent Services", is part of the "Service Modeling" topics in SOAMED. Consequently, it is related to the topics of Youssef Arbach and Cem Sürücü.

Cem Sürücü is concerned with adaptive process modeling. There already exists related work where effort was put into scenario-based modeling of adaptive processes; therefore, there are probably some similarities between these two topics. It is part of our future activities to work out where some collaboration could happen.

Youssef Arbach is concerned with graphical modeling of dynamic coalitions. As almost all scenario-based modeling techniques use graphical notations, there may be some points of collaboration. As for Cem Sürücü, this has to be worked out in the future.

Further, as the technique developed by Robert Prüfer is based on Petri nets, there are similarities to the topics of Richard Müller and Kristian Duske, who are also working on Petri net related topics. Consequently, all of these three Ph.D. students benefit from the knowledge and insight the others gain working with Petri nets.

2.5 Cem Sürücü

The process modeling of the medical treatment in the case studies mentioned below is rigid up to now and has to be adapted to real-life, dynamic terms and conditions of the individual treatment. The aim of Cem Sürücü's graduate thesis is to analyze and to extend existing methods of information technology, which can be utilized for the individualization of treatment processes and to achieve the ability for chronological and content-sensitive dynamic sampling as well as adaption to changing influences for a successful process modeling.

In this context Cem Sürücü uses case studies in the fields of interdisciplinary stroke and rheumatoid arthritis treatment. Nadim Sarrouh also examines acute stroke treatment and stroke-rehabilitation. Therefore, there is current communication on the case study and prospects of cooperation. There might be the opportunity to integrate Nadim Sarrough's formal modeling and Cem Sürücü's adaptive process modeling intention.

Further coherence exists to the work of Denny Schneeweiss, who develops a concept for a dynamic time- and resource planning system which allows adaptive re-planning during the treatment process. This depicts another opportunity to expand the scope of work within Cem Sürücü's thesis.

2.6 Youssef Arbach

2.6.1 Interrelations in Topics

I am working also on the same idea, roughly, of Nadim Sarrouh which is Dynamic Coalitions. It is generally the same topic but there are differences in orientation. I focus on finding graphical models for the problems of DC (Dynamic Coalitions) like dynamic membership and information flow. On the other hand N. Sarrouh focuses more on the privacy topic and access rights in DCs. Anyway, the difference will get bigger over time, but for the time being I am depending on his examples, papers and results. Later, we will try to publish at least one paper including our shared work together in this field.

On the other hand, my work overlaps with other's work forming another cluster. It is the cluster of modeling, where Cem Sürücü, Robert Prüfer and I are working to define different kinds of models for our different topics. Some of these models might be graphical like mine, and other might be mathematical and so on. Some of these models on the other hand will be oriented for the user to use, while others will be for the developers and specialists to deal with.

2.6.2 Integration With SOA

Both clusters: Dynamic Coalitions and Modeling lead in one way or another to the main topic of SOA (Software Oriented Architecture). The topic of DCs is related to SOA in a way that services and their combinations can be considered as a dynamic coalition of services in order to achieve a certain goal. On the other hand, looking at SOA from the perspective of being an Architecture, modeling is an essential goal and a necessary tool for users, designers and researchers.

2.6.3 Supervising

N. Sarrouh isalready helping me, and orienting me in some details. He is providing me some papers to read, references, in addition to providing support and motivation. Thanks to him, I hope we continue cooperating together.

2.7 Johannes Starlinger

Johannes Starlingers research targets the establishment of similarity measures for Scientific Workflows. The research area of Scientific Workflows is shared with Marc Bux, a Ph.D. student from the second generation of SOAMED students, who is exploring parallelization and distributed execution of these workflows. The interlink between both topics not only resides in the common object of study. It is also reflected in the application of Scientific Workflows to any problem requiring large scale computation. As a

consequence, Marc and Johannes are co-supervising a seminar held during the summer term 2012 which deals with the problem of large scale data analysis[1].

Closely related to Scientific Workflows are the fields Business Workflows and Business Process Models, investigated by SOAMED student Andreas Rogge-Solti. Similarity measures for Business Workflows have seen great scientific interest already and the methods developed for them provide a good starting ground for Johannes Starlingers endeavours.

2.8 Marc Bux

Marc Bux investigates parallelization techniques for scientific workflows from the field of next-generation sequencing. Scientific workflows are high-level compositions of sequential and concurrent data processing tasks. Increasingly large amounts of data produced in most fields of scientific research have led to a demand for strategies towards distributing workflow execution among several compute resources.

This research topic is most closely related to the work of Steffen Zeuch and Johannes Starlinger, both of which are also doctorate students in SOAMED. Steffen Zeuch researches parallelization from a multicore perspective, which shares many fundamental ideas with parallelization of scientific workflows. On the other hand, Johannes Starlinger focuses on similarity measures for scientific workflows, which overlaps with Marc Bux's topic in the underlying computing paradigm of scientific workflows. Together with Prof. Dr. Ulf Leser and Astrid Rheinländer, Johannes Starlinger and Marc Bux are also coordinating a seminar on large scale data analysis.

Beyond aforementioned intersections in research topics, Marc Bux also co-organizes a GMDS / GI interdisciplinary workshop together with SOAMED doctorate students Richard Müller and Christian Neuhaus. The title of this workshop is "Service-oriented Architectures in the Healthcare Domain".

2.9 Steffen Zeuch

Steffen Zeuch's research deals with Multicore and the Map/Reduce paradigm. The working title is *Bringing Map/Reduce and Multicore together*. Both technologies try to further speedup the computation of big datasets by parallelizing them on diffent units. The units can be cores of the same processor or different physical separated computers in a network.

This research area is most likely shared with the approach of Marc Bux, another Ph.D. student from the second generation of SOAMED, who deals with parallelization from the workflow point of view.

Furthermore this research task will interfere with Daniel Janusz within the new founded *Mobility Lab* in the database research group. The target of this lab is to evaluate the

[1] Please see https://www.informatik.hu-berlin.de/forschung/gebiete/wbi/teaching/archive/ss12/se_largescale/

SOA technologies for the mobile area. Because modern cellphones consist of multiple processors and cores, this will be another area of application for Multicore technologies.

2.10 Andreas Rogge-Solti

Andreas Rogge-Solti is working on business process intelligence methods in the health care domain with a focus on process monitoring in non-automated environments. His work deals with different questions that arise, when not all state changes are reported to the monitoring system. An example question is, where is it most useful to manually measure the progress of an instance. A follow up on that is: How many are necessary at all?

First collaborations inside SOAMED were achieved when working together with Richard Müller (Sec. 2.11) elicitating hospital processes at Charité Sozialpädiatrisches Zentrum. Key insights of this project were published at ZEUS 2011 in Karlsruhe [2]. In this work, particular requirements of hospitals were addressed, e.g. collaboration of multiple roles, and how the BPMN 2.0 language could be enriched to allow for support of these constructs.

Currently, A. Rogge-Solti is discussing ideas with Denny Schneeweiss (Sec. 2.12, on how to integrate the probabilistic estimation results with the constraint based planning and scheduling mechanism. We hope to achieve an improvement over traditional scheduling algorithms by the use of probabilistic estimations.

2.11 Richard Müller

Checking conformance of a given implementation to its also given specification is *the* central scientific problem of Richard Müller's thesis. As conformance checking is a rather general question, cooperations with other SOAMED Ph.D. students naturally arise.

Conformance checking is one of the most important types of process mining, which is settled in the workflow area. Thus, Richard Müller's work is closely related to the research of Andreas Rogge-Solti, and Denny Schneeweiss. In 2011, Andreas Rogge-Solti and Richard Müllers investigated healthcare processes at the Charité Sozialpädiatrisches Zentrum Berlin[1]. Their findings were published at ZEUS 2011 [2].

The research of Richard Müller focuses on formal verification and formal models, which is related to the research of Kristian Duske and Daniel Stöhr. Conformance checking on formal models of services has been formalized as substitutability. Since functional correctness - the research topic of Kristian Duske - and substitutability are closely related, there are many opportunities for future cooperation. Kristian Duske and Richard Müller surveyed formal models for timed services, which has been published at ZEUS 2012 [1].

[1] `http://spz.charite.de/`

Finally, Daniel Janusz and Richard Müller are supervising a seminar on SOA held during the summer term 2012[1].

2.12 Denny Schneeweiss

Denny Schneeweiss' research focuses on Constraint based Planning and Scheduling and is set within the Workflow-area of SOAMED, which is shared with Richard Müller and Andreas Rogge-Solti.

The goal of D. Schneeweiss' work is the development of a concept for dynamic, adaptive time- and resource-planning of treatment processes in medical facilities. Treatments required by individual patients as well as treatment-related activities like preparation- and cleaning-procedures can be represented by specific process models (e.g. in BPMN). These can be transformed into a Constraint-Satisfaction-Problem for which a constraint-solver can generate a plan that optimally allocates the process-activities to the medical staff and resources available.

This topic is connected to the work of Andreas Rogge-Solti, who focusses on the probabilistic estimation of the progress of processes in medical environments, where a process monitoring system can only gather sparse information that is mainly produced by the medical staff. The predictions generated by A. Rogge-Solti's model could serve as an input for the adaptive time-planning method developed by D. Schneeweiss to enhance the generated plans.

A further connection exists to the research of Cem Sürücü, who works on a methodology to adapt treatment processes to the needs and specific medical conditions of individual patients. This methodology could also benefit from enhanced time- and resource-planning.

2.13 Daniel Janusz

Daniel Janusz is working on *privacy-preserving query execution in medical workflows.* If a new medical workflow at hospital involves data exchange or automatic data processing, the workflow must to be approved by data protection officers. In general, privacy concerns arise whenever personal data is exchanged across company borders. This approach aims to handle these concerns by generating alternative privacy-preserving workflows. Thereby, a protocol was developed, that enables health care provider to query remote data sources about a specific patient.

The work of Christian Neuhaus and Daniel Janusz complement one another, as the protocol developed by Daniel Janusz can run on top of Christian Neuhaus' cloud infrastructure. Another cooperation within SOAMED currently starts with Steffen Zeuch in the context of a new *Mobility Lab* at Humboldt University. With the Mobility Lab we will simulate

[1] http://www.dbis.informatik.hu-berlin.de/lehre/semesteruebersicht/sommersemester-2012/soa-service-orientierte-architekturen-grundlagen-und-anwendung.html

ambient assisted living scenarios, which includes automatic privacy-preserving processing of so-called big data.

In 2012, Daniel Janusz and Kristian Duske are organizers of the annual meeting of the DFG research training groups that will take place at Schloss Dagstuhl[1]. Moreover, Daniel Janusz and Richard Müller are supervising a seminar on SOA held during the summer term 2012[2].

2.14 Christian Neuhaus

The topic of Christian Neuhaus' dissertation is *Exchange and Processing of Patient Data on Cloud Infrastructures*. The motivation of this topic is the current shift from local to cloud resources for computation and storage need. As medical data is highly sensitive with regard to data privacy, using cloud resources poses an challenge. The aim of this dissertation is to identify reusable design patterns for recurring problems in the design of software systems for processing patient data on cloud infrastructure. These patterns exploit the strategy of scattering and distributing power and responsibility among different service nodes to alleviate privacy and security concerns.

The work of this thesis touches the work of Nadim Sarrouh, who is regarding access control in distributed scenarios from a formal perspective and Daniel Janusz, who develops privacy-preserving communication protocols.

2.15 Nadim Sarrouh

Nadim Sarrouh's research deals with the creation of a formal modeling framwork for privacy-sensitive dynamic coalitions. Since privacy is an imporantant topic in SOA-development as well as in medicine, there are various researchers investigating this field from different perspectives, such as Christian Neuhaus and Daniel Janusz.

The closest coherence is to be found with Youssef Arbach, who works on a graphical modeling tools for dynamic coaliton scenarios.

One of N. Sarrouh current case studies in the medical field, is the rehabilation process for stroke patients at Charité Berlin. Cem Sürücü is also working at that case and therefor in close exchange with N. Sarrouh. There have been several talks to intgrate the formal modeling apporach of N. Sarrouh with C.Sürücü's efforts to create a machine learning tool in the ambulant treatment sector, which can be seen as a dynamic coalition too. Further possibilites or cooperation are currently investigated.

[1] please see http://dagstuhl2012.soamed.de/ [2] please see http://www.dbis.informatik.hu-berlin.de/lehre/semesteruebersicht/sommersemester-2012/soa-service-orientierte-architekturen-grundlagen-und-anwendung.html

References

[1] Duske, K., Müller, R.: A survey on approaches for timed services. In: Proceedings of the 4th Central-European Workshop on Services and their Composition (ZEUS) (2012)

[2] Müller, R., Rogge-Solti, A.: BPMN for Healthcare Processes. In: Proceedings of the 3rd Central-European Workshop on Services and their Composition (ZEUS) (2011)