WASP Software Engineering Course Module Assignment 2025

Antonia Welzel
Chalmers University

1. Introduction

My PhD project focuses on trustworthy resolution of requirement conflicts at runtime. The system domain for this research is mainly autonomous systems, and the aim is to semi-automatically resolve requirement conflicts with only the necessary amount of human involvement.

Requirement conflicts arise when two or more requirements are contradictory or incompatible and therefore cannot all be implemented successfully together at the same time. An example of this is for instance the time for the system to perform a certain task should be under 5 seconds, however this is not possible while keeping the daily costs of maintaining the system to under $1000. There are some approaches to modelling requirement conflicts as well as resolving them such as prioritisation, however they are not always very dynamic and able to consider the complexities of their environment, such as multiple stakeholders and goals being present. As a result, the existing solutions are often not very effective at runtime. This becomes further an issue in autonomous systems, which are designed to adapt to uncertainties in the environment particularly at runtime.

Moreover, with more autonomous and intelligent software solutions, the human's role changes and the control they have over the system, where the lack of control and more black-box nature of AI systems tend to negatively impact human trust. Yet, humans still need to trust the system and its ability to resolve conflict autonomously and therefore providing assurance or explanations that can promote the user's trust is another important aspect to resolving conflicting requirements at runtime in an effective manner.

2. Lecture Principles

One topic that was discussed in the lectures and is related to my Phd topic was non-functional requirements (NFRs) for AI-enabled systems. There is a strong focus on NFRs in AI-based systems, and new approaches for specifying requirements have been proposed to manage these systems as they involve a high level of uncertainty (Martinez-Fernandez et al., 2022). This would also in turn impact the conflicts that may arise in the system, where more effective specifications may prevent certain conflicts. Overall however many of the requirements engineering concepts and techniques we discussed appear to be more design time focused. There seems to be little focus on how to manage requirements, particularly conflicts like in my project, at runtime, which given the uncertainty that typically affects AI-enabled systems would make it an important topic to study further.

Another topic that was discussed in the lectures was quality assurance for software and more specifically software testing. While testing is not directly part of my project, it is an essential step that I need to consider when developing solutions to manage runtime conflicts. For the system to autonomously manage conflicting requirements at runtime, the uncertainty of the environment and the system's behaviour after deployment need to be taken into consideration. This presents many potential outcomes that are difficult to anticipate beforehand and also creates challenges for me as the developer to ensure proper testing and quality assurance at runtime so that in a proposed solution, conflicts are resolved as intended.

3. Guest-Lecture Principles

One guest lecture covered different concepts within requirements engineering. One topic that I feel is strongly related to my research is for instance goal modelling and goal refinement which can help with identifying conflicts. Conflicts between goals or requirements is the problem that I am trying to address in my project, and specifically finding solutions that would support the resolution of these conflicts at runtime. In order to resolve conflicting requirements at runtime, goals would need to be modelled in a dynamic way so that the system can handle new goals that may arise or changes in existing ones, and thereby manage potential conflicts between all these goals.

In the guest lecture from SAAB, we discussed the human aspects within software engineering and how human's trust in AI systems can often vary, where many do not trust it. While the human factors that come up in my project are generally more related to users rather than developers, one aspect that I think relates to my project is the trust aspect and how important it has become to consider when designing any system or solution that is connected to AI. So based on the lecture, one thing I will take with me in my project is to also think more about how to communicate and manage the risks that requirement conflicts present to its users and in turn help maintain their trust in the system.

4. Data Scientists vs Software Engineers

4.1. Do you agree on the essential differences between data scientists and software engineers put forward in these chapters? Why or why not?

I think the definitions and differences between the roles can sometimes depend on the specific context, however on a general level I would agree with the points discussed in the book such as the product, user, and business focus in software engineers versus data scientists' focus on model build and data. The book also discusses the importance of T-shaped individuals, especially for interdisciplinary teams to work well together, which I think is reasonable since both software engineers and data scientists would need a general understanding of the other team members' competence and level of understanding on different topics so they can effectively communicate and work together.

4.2. Do you think these roles will evolve and specialise further or that "both sides" will need to learn many of the skills of "the other side" and that the roles somehow will merge? Explain your reasoning.

On a very general level, I think in the future being a T-shaped individual will become more important. The roles may also become more merged as AI-based systems are becoming more prevalent and 'pure' software engineering may not be enough to develop effective systems, while AI engineers also need to be able to consider the entire system in their work and therefore need to integrate techniques and approaches from software engineering. However, on a more specific level, how these roles may change and what skills are needed from each side will probably depend on the type of team you are working in, such as what skills other team members have and how they can complement each other, as well as the ways of working and the organisation you are operating in, such as cultural factors or what products and industry you are working in, as some companies may value and invest more in one side.

5. Paper Analysis

5.1. Paper 1: "Towards a Roadmap on Software Engineering for Responsible AI" by Lu et al. (2022)

5.1.1. Core Ideas and SE Importance

The core ideas of this paper are about the software engineering aspects of developing and maintaining responsible AI systems. The authors present a roadmap on software engineering practices that promote reliable AI systems, which conceptualises how software engineering can help to engineer trustworthy AI systems and therefore provides important insights for both software engineers as well as AI engineers. The paper considers three different levels of governance for the software development and management aspects of responsible AI systems and discusses the current state and research challenges of each level to provide insights into potential guidelines or tools to support AI system stakeholders.

5.1.2. Relation to my Research

My thesis project focuses on managing conflicts and subsequent trade-offs between requirements at runtime while considering the impact of the user's trust. Different practices and solutions are discussed in this paper regarding requirements engineering for responsible AI systems such as the need for trade-offs between ethical requirements with the help of design patterns. However, the mentioned requirements engineering techniques, as well as the techniques for other software engineering phases in the paper, are relatively design time focused and therefore are only partly relevant to my project. While requirement trade-off considerations at design time are crucial for a system to also be effective and responsible at runtime, it does not offer support when new or unexpected requirements arise that you would want an AI system to handle in an autonomous manner. Nonetheless, I think the core ideas of the strategies that are mentioned in this paper, particularly in relation to requirements engineering, could be very useful for me in the future when designing an approach for runtime resolution of conflicting requirements.

Moreover, one of the key concepts of the paper is trust and trustworthiness. This is also one aspect I focus on in my project. In the paper, the authors present an overview of architectural patterns for responsible AI systems, which provides interesting insights into the human factors related to different design patterns as well as use cases for each pattern. While these ideas are again more design time activities, I think it is still important to think about when designing a solution framework or approach for runtime conflict resolution that is able to inspire and/or maintain trust in the system users by for example considering factors such as human control or how human values are addressed.

5.1.3. Integration into Larger AI-Intensive Project

In this scenario, I consider a smart home system that connects different devices to for instance manage the temperature, lighting, and the house's security, for example through cameras. The insights from the paper can be applied in the design of the system to ensure that it is reliable. For example, the architecture of this example system would require consideration of both AI and non-AI components that may exist in the network of devices to enable high integration between all components. Moreover, including features such as AI mode switchers or kill switches to enable the human to manage the AI capabilities of the system or override it would be critical for the type of system considered here.

How the system might handle conflicts that arise between either different user preferences or between the different devices' requirements would be something that I consider in my research. For instance, how to

handle a conflict between two system rules that lead to conflicting temperature settings so that the users feel satisfied and confident in the system. One example resolution could be letting the human decide which rule to prioritise, which would need to be considered during the design of the system's architecture and also relates back to the ideas in the paper.

5.1.4. Adaptation of my Research

Some of the research challenges that the authors raise in the paper are for instance in regards to trade-offs between ethical requirements. Handling trade-offs that involve ethical requirements in a balanced way is identified as an ongoing challenge since solving these conflicts often requires overriding one of the requirements. While I don't specifically focus on ethical requirements, conflicts involving them can often be difficult to resolve, especially at runtime since they might represent larger system issues requiring consideration of different stakeholders. However, they are often critical requirements and if I extended my project to also focus on managing specific cases that involve these types of conflicts at runtime, it could provide important results.

5.2. Paper 2: "Novel Contract-based Runtime Explainability Framework for End-to-End Ensemble Machine Learning Serving" by Nguyen et al. (2024)

5.2.1. Core Ideas and SE Importance

In "Novel Contract-based Runtime Explainability Framework for End-to-End Ensemble Machine Learning Serving", the authors focus on explainability at runtime in an AI system. They propose a framework to evaluate different metrics in ML contracts which would help improve the explainability in end-to-end ensemble ML serving at runtime. The framework includes an explainability constraint schema to evaluate various quality metrics in different conditions. Explainability at runtime is maintained through an observation agent that detects contract violations as well as issues in ML quality metrics such as performance issues at runtime. This paper plays an important part in the engineering of AI-enabled systems as it proposes a runtime solution for assuring the system's quality in a transparent manner, which provides useful insights for effective quality assurance in AI systems and would help support the system's users and potentially operators who maintain it after deployment.

5.2.2. Relation to my Research

This paper provides a framework for a transparent runtime evaluation of the system's quality. Providing runtime explainability of a system end-to-end is complex but crucial for providing a more holistic understanding of the system quality. While the quality evaluation aspects are not directly related to my research, the runtime explainability aspects of the paper's contribution are strongly related to what I want to achieve for a trustworthy conflict resolution framework, where explainability and especially explainability at runtime can help contribute to more trust. Moreover, as the authors state in the paper, future work includes further development of runtime analyses to achieve different optimisation goals. This could also involve evaluating different trade-offs between the goals, which would be strongly related to the aims of my research project.

5.2.3. Integration into Larger AI-Intensive Project

In this scenario, I again consider a smart home system with the same capabilities as in 5.1.3. ML contracts would be defined that include thresholds for the data quality, accuracy of model outputs and confidence levels. The observation agent monitors the system and its performance. Furthermore, explanation reports

are provided that describe how the constraints are met or not. This could then help an operator to keep track of the system's performance after deployment, potentially seeing that at certain times of the day the quality demands in the contract cannot be met and need to be adjusted.

When the system identifies and explains a quality issue in the ML system that does not meet the explainability constraints, the observation agent might propose a change in the contract to enable a better quality in the system. However, conflicts might arise between for example the need for higher quality data from the system's camera or sensors and energy consumption or cost constraints. In that case an approach for resolving these conflicts at runtime, like I'm aiming for in my project, could be used in the system to mitigate the issue.

5.2.4. Adaptation of my Research

The challenges that Nguyen et al. (2024) are aiming to address with their contribution within runtime explainability are the system's ability to provide insights into the system's quality and resource usage at runtime rather than making the system more interpretable for developers to support the system design, which has been the main focus in past research. The authors further identify challenges such as adequate real-time end-to-end monitoring and analysis, the modelling of certain ML-specific metrics or the need for continuous feedback between the user and system so runtime data can be collected and used for runtime explainability, which is not widely supported in existing ML models.

My research is currently not directly connected to these challenges, since it explores how a human should be included in the system's decision-making when resolving conflicting requirements to enable trust as well as effective resolution, since some tasks require human insights. As mentioned above, one key aspect of trust is transparency and explanations which need to be provided at runtime in a way that is accessible to humans. For the system to have awareness of when and what explanations it should provide to a user as well as when to involve them in the decision-making, also in regards to the human's trust, it needs certain monitoring and analysis capabilities similar to what is mentioned in the paper here. Therefore, research into achieving these types of system capabilities would be an important aspect to also include in my Phd.

6. Research Ethics and Synthesis Reflection

6.1. Search and Screening Process

In the process of finding the two research papers discussed in the previous section, I searched both in the ACM database and through Google scholar for full research papers in the CAIN proceedings that relate to my research area. I used different combinations of search terms to find papers that relate strongest to my research area. These terms were 'conflict*', 'trade off', 'runtime' and 'trust*'.

6.2. Pitfalls and Mitigations

When searching for the literature in both the ACM database and Google Scholar, I tried first the terms 'conflict*', 'runtime' and 'trade off' since they relate strongest to my project. I didn't encounter any misleading titles or abstracts, however when I looked through the abstracts of the papers that came up, no papers explicitly focused on requirement conflicts or they mainly focused on managing trade-offs at design time. Therefore, I added 'trust*' to my search string to find results that are more relevant to my project, due to the lack of results directly related to requirement conflicts at runtime. For my final

selection, I chose the papers that I found to be best related to trust and software engineering of AI with at least some consideration of runtime activities.

.

6.3. Ethical Considerations

The steps I took to ensure originality in my paper selection was to look at their method sections and acknowledgements to see how the research was conducted and how AIs such as LLMs may have been used. The steps I took to ensure originality for my discussion of the chosen papers was to cite any sources that I am referencing and not use any LLM for writing my reflections.

References

Lu, Q., Zhu, L., Xu, X., Whittle, J., & Xing, Z. (2022). Towards a roadmap on software engineering for responsible AI. *In Proceedings of the 1st International Conference on AI Engineering-Software Engineering for AI*. pp. 101-112.

Martínez-Fernández, S., Bogner, J., Franch, X., Oriol, M., Siebert, J., Trendowicz, A., Vollmer, A. & Wagner, S. (2022). Software engineering for AI-based systems: A survey. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 31(2), pp. 1-59.

Nguyen, M. T., Truong, H. L., & Truong-Huu, T. (2024). Novel contract-based runtime explainability framework for end-to-end ensemble machine learning serving. *In Proceedings of the 3rd International Conference on AI Engineering-Software Engineering for AI*. pp. 234-244.