Software Engineering Assignment

Ayesha.jena

Begin with an introduction/abstract to your research and topic area. This should be a maximum of 400 words and should provide (just) enough basis so that your answers to the questions below can be understood.

In Industry 5.0, robots will interactively work together with humans in creating a shared workspace environment. Such interactions would not only be applied in structured or industrial-type settings but also in collaborative and unstructured environments. The main challenge here would be developing robots that can handle these complex settings. My project focuses on how robots and humans could work in collaboration within such unstructured scenarios. One such way of effectively addressing this is Mixed- initiative interaction. Mixed-initiative interaction in Human-Robot Interaction(HRI) focuses on effective strategies for developing an interaction and reasoning framework to facilitate optimal collaboration. This system needs a deep understanding of the whole situation to be addressed including the intentions of the behind observable human actions. The framework aims to classify various types of intentions, such as initiation, movement, retrieval, idle observation, and composite intentions. Such classification is also essential for addressing one of the main challenges in HRI, i.e. bi-directional communication. Thus, making it essential for shaping the robot's behaviour to ensure effective communication.

As part of the work, we utilized Mixed Reality systems to develop a framework for interaction. This let us understand human interventions in various scenarios. Thus, providing valuable insights into interaction patterns. With this, we will be able to refine our understanding of how humans and robots can collaborate better, ultimately leading to safer and more intuitive interactions in real-world environments. Furthermore, intention reading comprises an understanding of the goals of observable actions, which is at the core of natural and safe human robot interaction.

Select at least 2 principles/ideas/concepts/techniques from Robert's lectures and discuss how they relate to your research and topic area.

Behavioral Software Engineering: In terms of software engineering part, one principle that I have tried implementing in the system designs is modular design, testing, and user-centered design. HRI focuses on humans as a participant, observers or users of the system. This necessitates the need to understand how to make the robotic systems usable in the presence of humans through interaction design for robots in social settings.

Testing and Quality Assurance: In my view, developing robotic systems is a further step in software engineering with the additional component of specialized hardware and thus having to have real-time processing, integration, testing, safety, and maintenance in long-term use cases. As mentioned in the lecture:

SE4ML = TradML + People + Process + Business + Modeling + QA + Maintenance + Ethics [3]

Similarly, HRI = Hardware (Sensors + Actuators + Embedded Systems) + Software (Control Systems + AI/ML Algorithms + Real-time Processing + Robot Operating System) + Human Factors + Integration + Testing + QA + Maintenance + Safety + Ethics among others. As Software engineering and AI engineering are an integral part of the system architecture for robotics and its application domain, quite some of the concepts would apply to the field of human-robot interaction with little or no modifications.

Select at least 2 principles/ideas/concepts/techniques from the guest lectures and discuss how

they relate to your research and topic area
As explained above, Behavioral Software Engineering is crucial for Human-Robot Interaction. It focuses on understanding, designing, and evaluating robotic systems planned to be used by or in collaboration with human beings. As mixed-initiative interaction inherently involves communication between robots and humans in a bidirectional and safe way, it means understanding how humans operate in collaborative scenarios for the robots to be able to take action in a meaningful manner.

During the SAAB lecture, there was a discussion about different levels of automation, a concept that is widely used in robotics. Mixed initiative interaction plays around this concept of variable autonomy but in a way which is suitable for both humans and robots to benefit from the scenarios. That means the variability has to be designed keeping in mind the current level of automation, safety of AI techniques, and ethical considerations.

Find two full/long papers published in one of the CAIN conferences (3 have been held so far, all linked from `https://conf.researchr.org/series/cain` to an external site.), download and read them and then write in your assignment, for each paper

Describe the core idea(s) of the paper and why it/they are important to the engineering of AI systems

Paper 1: What About the Data? A Mapping Study on Data Engineering for AI Systems

The paper explains a mapping study on data engineering for AI systems called AI data engineering. They fully support the fact that our attention should be given to quality data and emphasis should be provided on that aspect beyond the preprocessing step [4]. In this paper, there are 4 central questions:

- (RQ1) Which data and AI engineering lifecycle phases are covered,
- (RQ2) Which technical solutions (tools/frameworks/platforms) for AI data engineering are proposed,
- (RQ3) Which architectures for AI data engineering are proposed, and
- (RQ4) What lessons are learned on AI data engineering? [4].

These helped in identifying the lack of data architecture guidances for AI systems. Further, the work also mentions that DataOps, which is similar to DevOps is an important focus area and the need for best practices and frameworks for data engineering part of AI engineering [4].

How the paper relates to your research:

As mentioned before robots operate in dynamic, unstructured environments, making effective data engineering critical to ensure accurate processing and interpretation of sensor data for real-time decision-making. In mixed-initiative systems, where tasks and authority are dynamically assigned, relying on data pipelines to integrate various data sources and ensure efficient processing of both human and robot inputs would be the effective method. For robots to understand and predict human intentions, they must process large volumes of data from the environment about semantic relationships, human motion, pose etc. The concepts of data-centric AI and lifecycle phases outlined in the paper could provide a structured approach to managing this data effectively. Additionally, the architectures and lessons learned from AI data engineering for developing best practices were also useful in understanding how could these be implemented in real-time systems.

How your research and its results would fit into a larger AI-intensive software project where one of the core ideas from the paper would benefit the project if applied. Describe both how the paper could help improve the project and how your WASP research would fit into the project.

The paper takes examples of good data engineering practices from different works and papers. Some useful findings that could enhance my project into a larger AI-intensive software project are as follows:

- Improving data quality and management through best practices like Breck et al [6, 4] who suggested a data validation system to detect anomalies for the data which is fed to the model and pipeline, would for example help with the gaze anomalies in my existing work [5] when a model is deployed for using this to navigate the robot in high-risk search and rescue environment.

- Derakshan et al. [2, 4] proposed continuous adaptation of models based on changes in incoming data. Such training shall be beneficial for my research, as it enhances the ability of robots to understand and estimate human intentions in dynamic environments. Integration of algorithms for continuous learning may result in the real-time adaptation behavior of the robot to be much more effective and efficient in the collaboration with humans.

Discuss briefly how your research could be potentially adapted/changed to make AI engineering in the project based on the idea of the paper even better/easier.

With the datacentric AI (DCAI) [4] approach, as suggested in this paper, high-quality data management and pre-processing will be well taken care of, which is one of the most important aspects for effective training of the AI-based model in a collaborative environment. Adoption of an end-to-end pipeline tool and framework that manage structured and unstructured data will provide consistency and quality to the data being dealt with during the flow of the project. Thus leading to reliable AI systems.

In addition to this, most of the experiments are currently done in controlled simulated environments to ensure control over external variables. Refining interaction designs and enhancing real-time data processing can optimize the collaboration process. Currently, the insights we gained from comparative studies [5] on human-assisted and system-assisted approaches, would help enhance mixed-reality interfaces that reduce the cognitive load and improve decision-making.

Real-time data processing frameworks for dynamic analysis of the interaction can immediately allow for variable autonomy adjustments and improvements to be sure that the AI systems would be user-friendly, efficient and effective in high-risk settings, and enable further improvement on the capability as well as the operational effectiveness of the human-robot teams.

Describe the core idea(s) of the paper and why it/they are important to the engineering of AI systems

Paper 2: Investigating the Impact of SOLID Design Principles on Machine Learning Code Understanding

This study explores the application of design principles, such as the SOLID principles [1], in a machine learning context by using controlled experiments. The goal of this study is to improve the readability, maintainability, and extensibility of ML code. The authors conducted a controlled experiment with data scientists and found that ML code designed according to SOLID principles is more comprehensible. They recommend integrating those software engineering best practices in the ML community so that in general, the code produced with ML is better, more maintainable, and even qualitatively superior.

Application of the SOLID design principles in machine learning (ML) [1] code is crucial for AI engineering, since it significantly increases clarity of code, ease of maintenance, and flexibility that are important in machine learning projects. These principles ensure that code is modular and easy to manage, thus allowing AI engineers to work with the code at ease without disrupting existing code. For example, extending functionalities and thus facilitating faster execution and deployment of newer models. A structured approach would benefit where there are complex algorithms or large datasets used as well for reducing redundant work which arises due to complexities in code management.

How the paper relates to your own research

The main idea of using the SOLID principles is the importance of using a structured approach when designing systems, and maintaining code. Similarly, my work is aimed towards working with an HRI interface which consists of multiple smaller subsystems being put together into one. There could be an

issue with maintaining the system over the long run [7] and performing version control and maintenance of every subsystem involved in it. However, I have tried to implement modularity in each subsection, not just by providing a structured approach but also a blueprint of the core implementation logic and its breakdown. Thus facilitating a blueprint to replicate it based on other software, and algorithms available but also making it usable in a wider context. This paper enhanced my understanding of this further which could be used to improve future iterations of the interface.

How your research and its results would fit into a larger AI-intensive software project where one of the core ideas from the paper would benefit the project if applied. Describe both how the paper could help improve the project and how your WASP research would fit into the project.
As discussed in the previous answer, integrating the findings from the SOLID principles paper would further enhance the modularity and reusability of project components reliably. That is sub-systems of the human-robot interaction interface are easier to manage, update, and expand. Even if new changes are needed, the system operates flexibly. Thus by combining my findings on effective human-robot collaboration, and HRI interfaces with a SOLID-based software framework, we can ensure that the collaborative aspect of the interface works fine. It would also lead to transparency during handover scenarios leading to more effective and harmonious outcomes or easy-to-detect errors in the system.

Discuss briefly how your research could be potentially adapted/changed to make AI engineering in the project based on the idea of the paper even better/easier.
I believe this could be done in a multitude of methods. As discussed before the HRI interfaces are made modular to some extent. However, developing clear protocols for human-robot interactions(variable autonomy handovers) that rely on real-time feedback and trust, can refine the robotic systems to be more intuitive. Additionally, modular components similar to a SOLID-compliant principle[1] could ensure additional system functionalities with less time. Thus, focusing on user-centred human-in-the-loop design principles, through experiments and iterative processes would also reduce mental effort and improve situational awareness from the user end, making it easier for people to work effectively with AI systems.

References

[1] Raphael Cabral et al. “Investigating the Impact of SOLID Design Principles on Machine Learning Code Understanding”. In: *Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering-Software Engineering for AI*. 2024, pp. 7–17.

[2] Behrouz Derakhshan et al. “Continuous Deployment of Machine Learning Pipelines.” In: *EDBT*. 2019, pp. 397–408.

[3] Robert Feldt. *Lecture notes in Software Engineering Cloud Computing course*. June 2024.

[4] Petra Heck. “What About the Data? A Mapping Study on Data Engineering for AI Systems”. In: *Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering-Software Engineering for AI*. 2024, pp. 43–52.

[5] Ayesha Jena and Elin A Topp. “Towards Understanding the Role of Humans in Collaborative Tasks”. In: *7th International Workshop on Virtual, Augmented, and Mixed-Reality for Human-Robot Interactions*. 2024.

[6] Neoklis Polyzotis et al. “Data validation for machine learning”. In: *Proceedings of machine learning and systems* 1 (2019), pp. 334–347.

[7] David Sculley et al. “Hidden technical debt in machine learning systems”. In: *Advances in neural information processing systems* 28 (2015).