Course: T520051101 - Interactive Systems Engineering (ISE), Fall 2023.
Exam Assignment

Prompt Engineering Interactive Systems with Large Language Models: An Experience Report

You will work **in groups** of 6 students throughout the semester and produce an **experience report**. The report should document the group work on **engineering an interactive system through conversations with ChatGPT**. You should put yourself in the end users' shoes, which means **your prompts to ChatGPT should be as close as possible to natural language conversations**. For that, the group must:

a) define (in common agreement with the teacher) an interactive system of their interest to be developed,
b) design task models, BDD scenarios, and GUI prototypes for such a system to clarify its requirements,
c) engineer such a system by **only** prompting ChatGPT to generate the expected source code. Direct source code writing (i.e., outside ChatGPT) **is not allowed**. If slightly fine-tuning the output code is necessary for whatever reason, this must be reported and justified.
d) engineer (also by **only** prompting ChatGPT) BDD GUI automated tests to verify whether the generated interactive system behaves as expected (i.e., according to the original requirements). Likewise, direct test code writing (i.e., outside ChatGPT) **is not allowed**. If slightly fine-tuning the output code is necessary for whatever reason, this must be reported and justified.

The final group report must be **in English** and **produced in LaTeX** using the ACM Primary Article Template (**max. 20 pages excluding references and appendixes**). An Overleaf version is available at: https://www.overleaf.com/latex/templates/acm-journals-primary-article-template/cpkjqttwbshg.

The report must include **7 sections** (see below). **All group members should contribute to all sections**. However, for individual evaluation purposes, **each group member** should be responsible for and **"sign" one of the sections**, **except the "discussion" section** to which all group members should **contribute equally** (therefore, no "signature" needed in this section).

- **Report Sections:**

1. Introduction (Should be "signed" by one group member)

In this section, you should introduce the topic of the report ("prompt engineering interactive systems with large language models") by explaining the **context** and **motivation** behind the study. This section should also clarify the **problem**, define the **objectives**, and briefly summarize the **approach** taken to solve the problem, as well as the key **results** and **contributions**.

2. Background and Related Work (Should be "signed" by one group member)

In this section, you should provide the **background knowledge** about the key areas involved in the study, namely *interactive systems engineering*, *end-user development*, and *generative AI for code generation*. In this section, you should also review prior **related work** about these topics.

3. Study Design (Should be "signed" by one group member)

In this section, you should describe how the study was organized and the **stepwise approach** taken by the group **to achieve the objectives** of the study. In this section, you should also define the following **metrics** to be observed and reported during the study: *number of interactions required with ChatGPT to produce the desired results*, *quality and precision of the answers received* (measured through e.g., the *number (and type) of errors introduced into the code, and the number (and type) of required*

fixes/fine-tunings), time spent in the development, etc. This study is an **experience report** and should be therefore designed and reported as such (refer to the course on scientific methods).

4. Case Description (Should be "signed" by one group member)

In this section, you should detail the **case selected for the study**, i.e., the interactive system that was agreed with the teacher to be developed. The section should explain the **main characteristics** of the chosen system, its **requirements**, **key business rules**, and **target users**. This section can make references to the designed task models, BDD scenarios, and GUI prototypes (to be included as appendixes).

5. Results (Should be "signed" by one group member)

In this section, you should present the **results** obtained by prompt engineering the interactive system in question. A **link to the GitHub repository** with the final source (and test) code of the generated interactive system should be provided here. This section must be divided into "**Source code generation**" and "**Test code generation**". The group should report on the **strategies employed to make ChatGPT generate appropriate code (and tests)** for the intended interactive system, and **how the multiple pieces of code generated by ChatGPT were glued together to build a coherent and executable system from back to front end**. Figures with examples can be used to illustrate inputs and outputs. Results regarding the metrics defined in the "Study Design" section should also be presented here. This section can make references to the appendix with the full conversation with ChatGPT.

6. Discussion (All group members should contribute **equally** to this section, therefore no "signature" needed)

This is the core section of the report and must be divided into two subsections. The first one should reflect on both the **effective and ineffective prompt strategies** adopted by the group and discuss the analyzed **metrics** (see the "Study Design" section). The second one should reflect and discuss on the **implications, challenges, and opportunities** regarding the following aspects of interactive systems engineering:

- **Usability and user experience**. For example,
    - *How may prompt engineering interactive systems with large language models affect the usability and user experience of the generated interactive systems?*
- **Interactive systems design processes**. For example,
    - *How does prompt engineering interactive systems with large language models affect the current user-centered design processes?*
- **User-centered requirements engineering and testing**. For example,
    - *How are user-centered requirements engineering techniques (such as personas and BDD scenarios) affected by prompt engineering interactive systems with large language models?*
    - *How can design decisions on GUI prototypes be incorporated when prompt engineering interactive systems with large language models?*
    - *How is user interface evaluation affected by prompt engineering interactive systems with large language models?*
    - *How is GUI automated testing affected by prompt engineering interactive systems with large language models?*
- **Model-based engineering of interactive systems**. For example,
    - *How is task analysis and modeling affected by prompt engineering interactive systems with large language models?*
    - *How can cognitive and human-error aspects be anticipated and taken into account when prompt engineering interactive systems with large language models?*
- **Consistency assurance between requirements and artefacts**. For example,

- *How to ensure the consistency between the user requirements and user interface design artefacts when prompt engineering interactive systems with large language models?*

Finally, for both subsections, the discussion should also include **how the results achieved by the group confirm and/or differ from prior work** (look back at the section "Background and Related Work").

7. Conclusion and Future Works (Should be "signed" by one group member)

In this section, you should **wrap up the report**, by recapping **whether the objectives of the study have been achieved**, **under which conditions**, and **what are the key findings**. This section should also point out **open questions and future works** that, in the group's view, would be required to further investigate new contexts/situations that weren't explored by the present study.

- **Final Submission:**

*(**Only the group leader** must submit the final hand-in to Digital Exam on behalf of the group.)*

The final submission to Digital Exam must include:

1. The **report** as a **single PDF** including: *(i)* the main text (20 pages); *(ii)* the list of references; and *(iii)* a set of appendixes with the designed task models, BDD scenarios, GUI prototypes, and the full conversation with ChatGPT.
2. The **full LaTeX project** as a **single ZIP file** including the source and .bib files, as well as all the figures and appendixes.
3. The full **source (and test) code** of the generated interactive system as a **single ZIP file**.

Additionally, the final **source (and test) code** of the generated interactive system **must be uploaded to GitHub** and the link to the repository provided in the "Results" section of the report.