

# Feedback Report (Week 06)

Student ID 2021315385

Name 이건

Title: Mobile GUI Agent

Speaker: Professor Sunjae Lee

After attending this week's seminar, please write a brief summary of the content along with your comments. Your response should be approximately half a page in length.

이번 주 세미나를 듣고, 세미나의 내용을 간단히 요약한 뒤, 이에 대한 본인의 생각이나 의견을 작성하시기 바랍니다. 전체 분량은 약 0.5페이지 내외로 작성하여 제출하시기 바랍니다.

## 1. Summary

Briefly summarize the main content of this week's seminar.

이번 세미나의 주요 내용을 간단히 요약하세요.

This week's seminar focused on mobile GUI agents also known as action-based agent. These action-based agents use Large Language Models (LLMs) to understand and operate UI screens and perform tasks requested by the user. However, limitations due to high operating costs and unreliability became apparent. The professor's research discussed in the seminar tackles these limitations. High operating costs and latency were solved by designing a human-like app memory that can learn tasks faster and cheaper as it accumulates more tasks by decomposing the tasks into subtasks, efficiently storing task executions, and flexibly recalling these task executions. As for the unreliability of these mobile agents, using a logic solver to validate the mobile agent's actions drastically improves the accuracy, error recovery and cost efficiency of the agent. The logic solver will translate the user's request into a domain specific language

using the combination of a translation model and recover model, which will then be used to verify the mobile agent.

## 2. Comments

Please include your comments, noting what you found interesting, what you would like to learn more about, and your personal thoughts on the seminar.

이번 세미나에서 흥미로웠던 부분, 더 알고 싶은 주제, 그리고 본인의 생각을 자유롭게 작성하세요.

It was interesting to learn about an action-based agent that sort of emulates the human's way of executing tasks. The charts surely prove that this way of learning tasks may be a slow-starter, but still has better results when it runs on a warm start. In the future, I would like to see a glimpse of the framework and coding of the logic solver, seeing how cost efficient and effective it is in validating responses of the mobile agent.