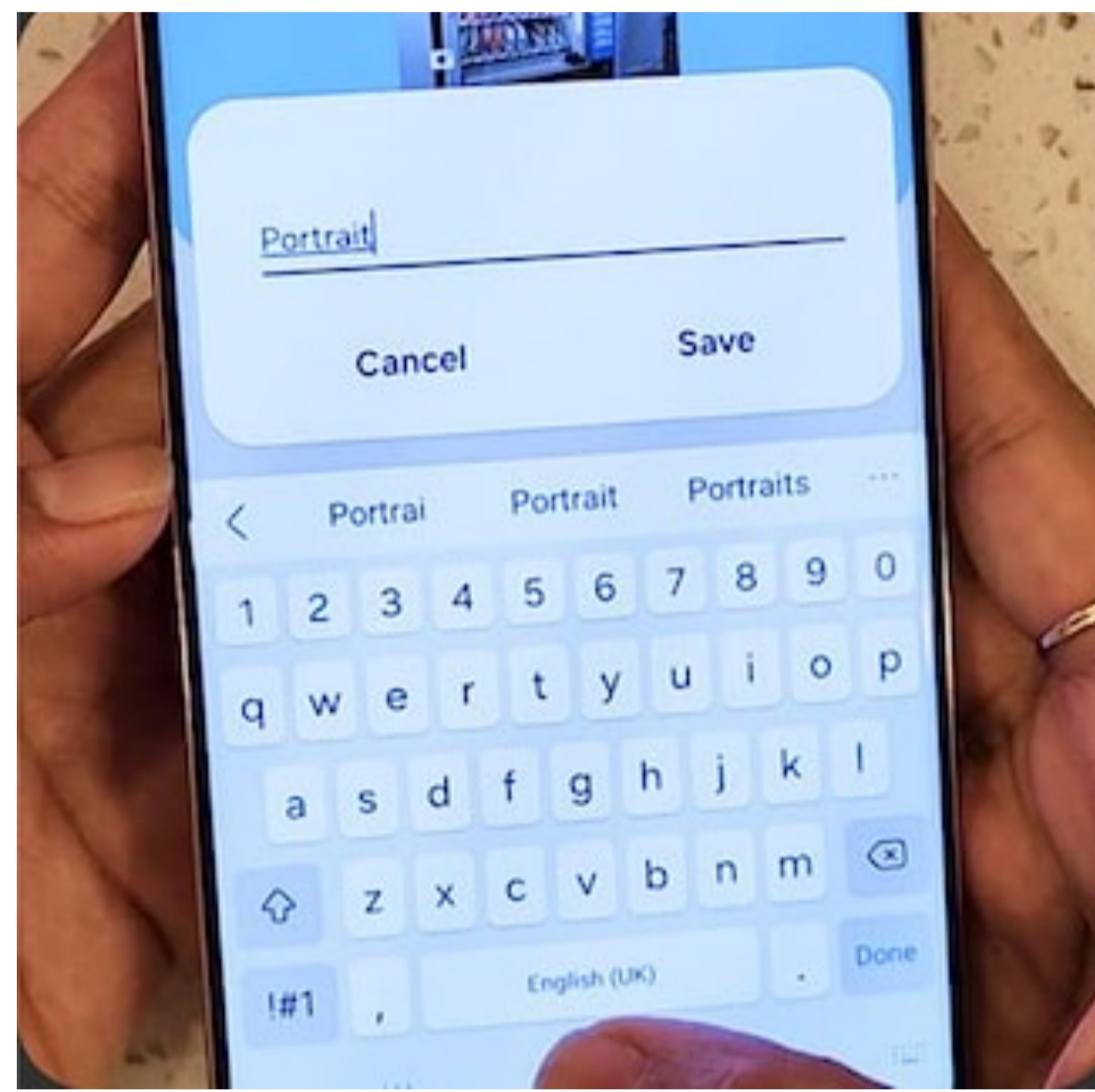# Poster: Efficient and Accurate Mobile Task Automation through Learning from Code

Shihe Wang, Li Zhang, Mengwei Xu

Beijing University of Posts and Telecommunications, Beijing, China
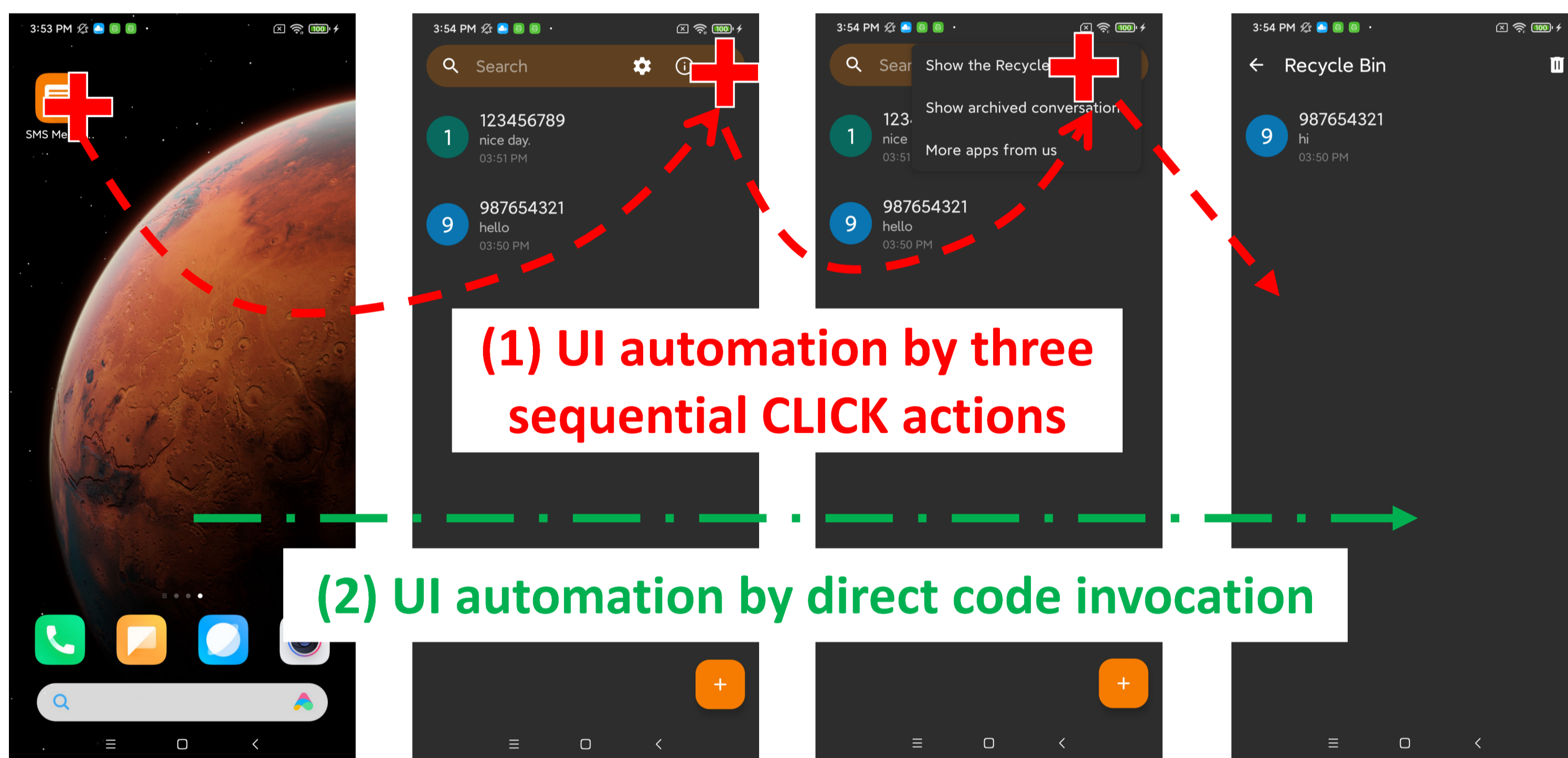
{shihewang,li.zhang,mwx}@bupt.edu.cn

Goal

How to empower mobile agents to automate daily tasks efficiently and accurately?

Mobile Task Automation through UI

Mobile Task Automation through code

(1) UI automation by three sequential CLICK actions

(2) UI automation by direct code invocation

A task "check my recycle bin in SMS Messages"

Introduction

➤ With the emergence and continuous prosperity of large language models (LLMs), mobile agents have experienced rapid advancements.

➤ Most LLM-based mobile agents merely imitate human operations. They **execute actions on graphics user interfaces (GUI)**, introducing two limitations: **inaccuracy** and **inefficiency**.

Problem

| Cause |

| A simple task requires multiple GUI operations! | Lengthy execution path on GUI is prone to error in individual steps. |

| What does this weird button mean? | Not all interactive information within the app is reflected in the GUI. |

| It takes too long to finish a simple task! | Feeding every GUI to LLMs leads to high latency and high computation overhead. |

Our Design: Learn from App Source Code and Write Code!

A three-layer architecture to make code execution as a supplementary to GUI operations.

➤ **Brain**: Generate fine-grained, step-by-step tasks from high-level task instructions.

➤ **Executor**: Generate code to execute tasks (e.g. activity switching), or execute directly operate on mobile GUI (e.g. click or swipe). A router is responsible for decision.

➤ **Interface**: A bridge for invoking generated code/UI actions through adb.
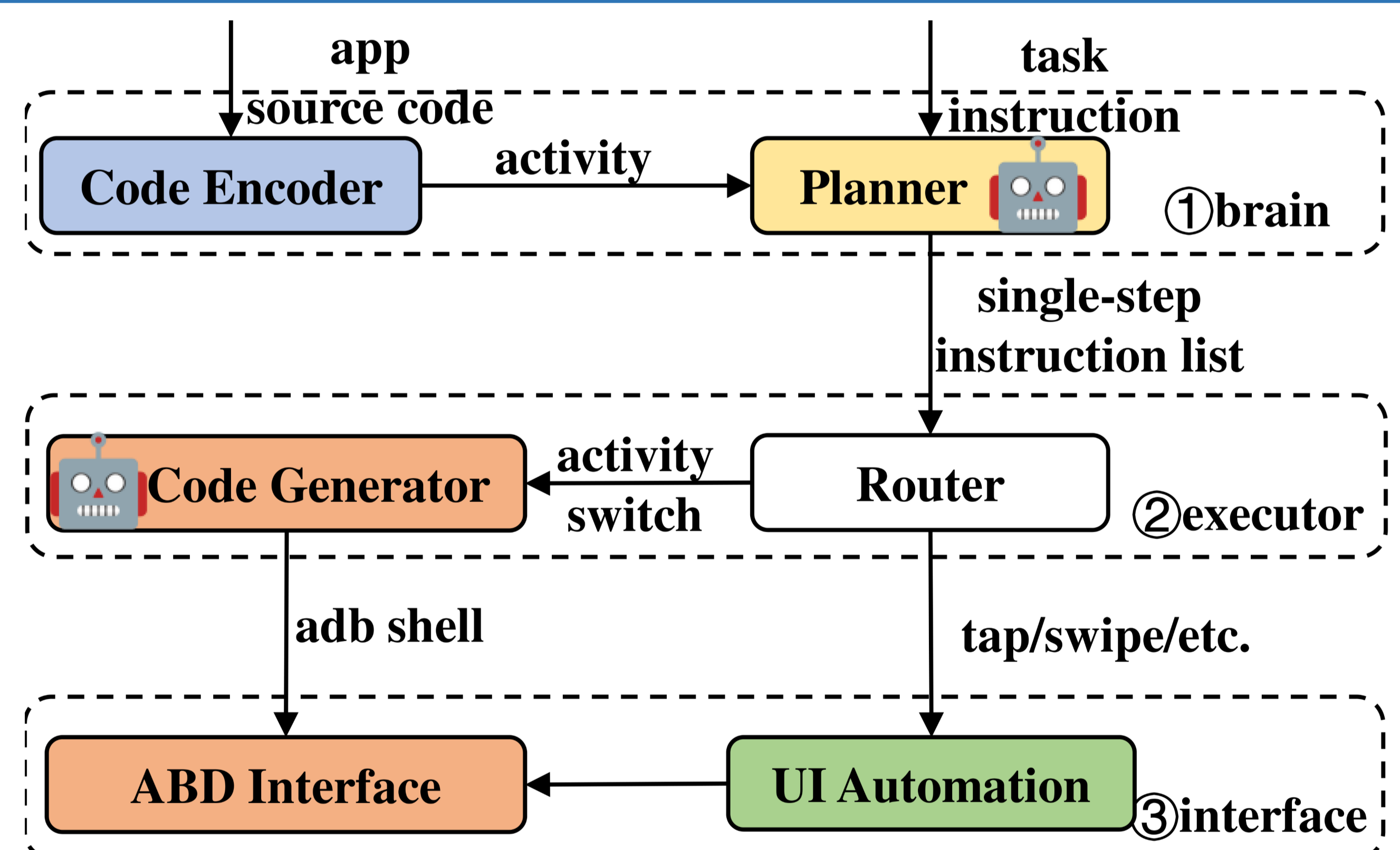
Preliminary Evaluation

Apps	Task Counts	Baseline Steps	Steps	TCR
SMS Message	5	12	5 (58%↓)	80%
Calendar	4	9	4 (55%↓)	50%
Gallery	5	14	5 (64%↓)	60%
Contacts	4	9	4 (55%↓)	50%
Notes	4	7	5 (28%↓)	50%
Total	22	51	23 (54%↓)	59%

*Baseline Steps: Number of total steps for GUI operations.
*Steps: Number of total steps. *TCR: Task completion rate.

➤ A dataset including 5 open-source Android apps and 22 tasks.
➤ Reduce the average number of steps to complete a task by up to 54%.
➤ Achieve an average task completion rate of 59%.

Follow the link to checkout our dataset: https://github.com/LlamaTouch/LlamaTouch /tree/main/dataset/SimpleTools