Error Type: Unclear Task
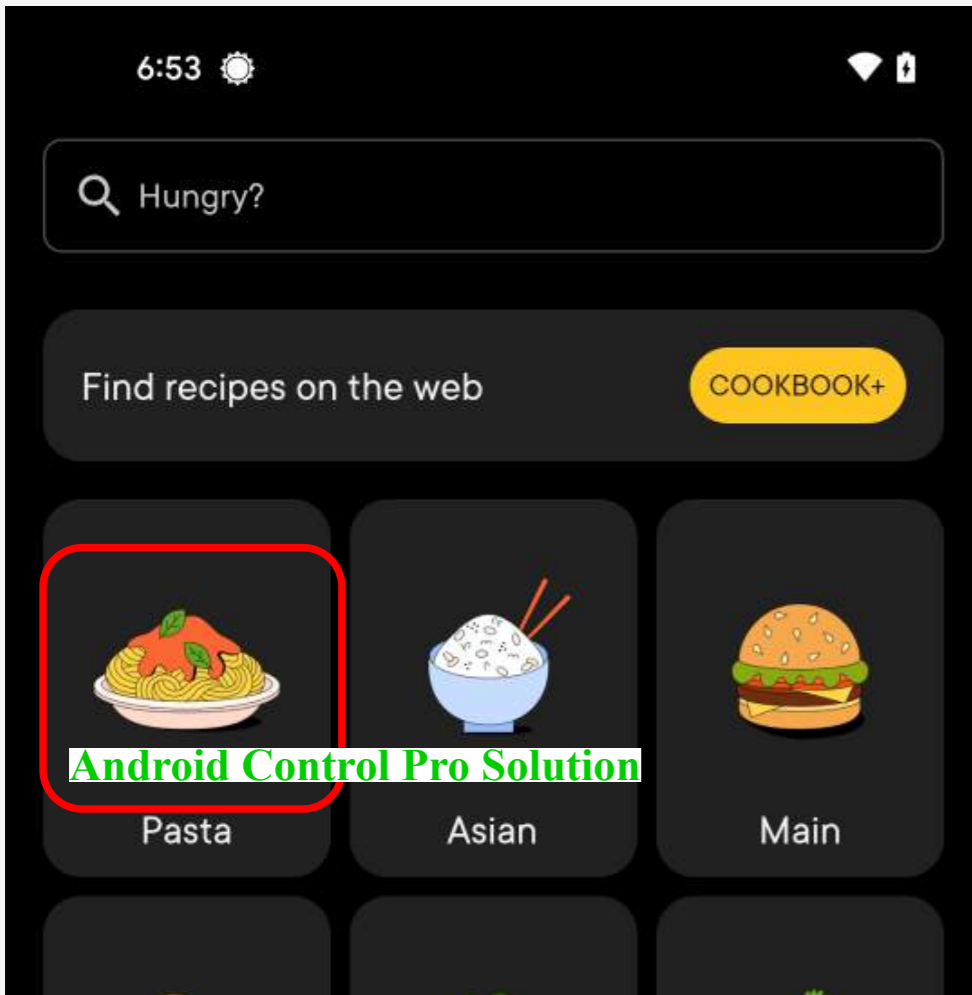
Android Control Problem:

Task: Tosso-dwi is the best literature artist. I get inspiration from the literature artwork created by Tosso-dwi in the DeviantArt app.
Memory: Step1: Click on the Search section;
Action Type: Click
Action Box: [15, 451, 1065, 1096]

Android Control Pro Solution:

Task: Explore the Daily Deviations section in the DeviantArt app to find artworks by Tosso-dwi.
Memory: Step1: Open the DeviantArt app; Step2: Navigate to the homepage.
Reason: The task description is vague and does not specify any action related to Tosso-dwi's artwork or the search functionality.
The memory only mentions clicking the search section, which is not aligned with the ground truth location.
The task lacks clarity and does not provide a clear instruction on what needs to be clicked, leading to confusion for the models.
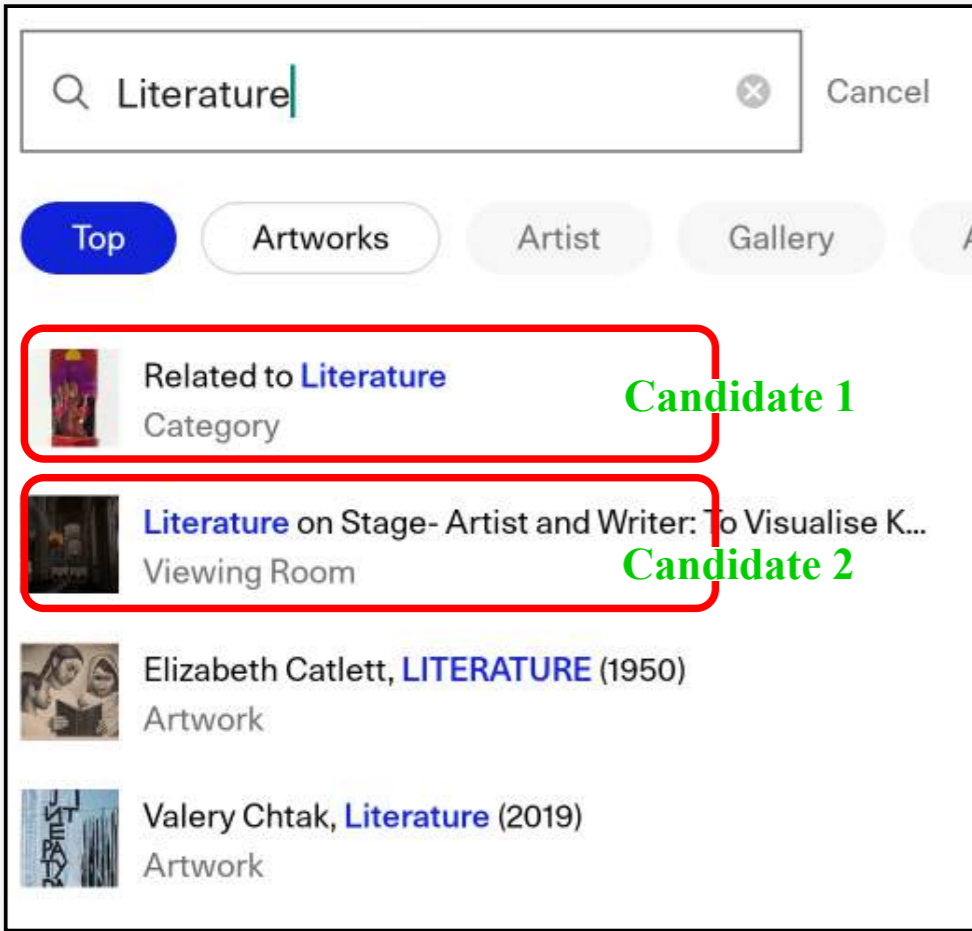
Error Type: Multiple Valid Actions

Android Control Problem:

Task: Look for a venison goulash in the kitchen stories app for me, I'm starving.
Memory: Step1: Click on the Search section;
Action Type: input_text:venison goulash
Action Box: None

Android Control Pro Solution:

Action Type: Click
Action Box: [80, 193, 916, 234]
Reason: The model predictions focus on clicking the search bar, which is a necessary step before inputting text, suggesting the ground truth should have included this action. Other categories are not applicable because the task and memory are clear, and the model logic is sound..

Error Type: Wrong Ground Truth

Android Control Problem:

Task: Literature art is my favorite art, so look for it on the Artsy app.
Memory: Step1: Open the Artsy app;Step2: Go to the search section;Step3: Search for Literature;
Action Type: Swipe: left
Action Box: None

Android Control Pro Solution:

Candidate 1: Action Type: Click Action Box: [80, 193, 916, 234]
Candidate 2: Action Type: Click Action Box: [80, 256, 916, 330]
Reason: the ground truth action of swiping left would not accomplish the task objective and appears to be mislabeled. The models that predicted click actions on the search results demonstrate correct understanding of the task flow, while the ground truth fails to provide actionable guidance for completing the stated objective.