

We record the inferred results for each input symbol. The result shows that LLM can correctly infer the click path for 97.22% of actions.

Among them, there are 7 (19.4%) actions that require manual provision of XPath information for a certain action to LLM in order to complete the inference. This is because during development, there is interference from component names, causing the LLM to incorrectly click on an unexpected component. For example, the action to be performed is to remove device, and there are two components on the interface. One is the “device details”, and the other contains the word “Edit”. Based solely on the text semantics, the LLM is more likely to choose “Edit” because it thinks this is related to editing the device and then removing it.

The reasons for failure include: some actions are complex to execute, since OOB (Out-of-Band) information is required. For example, when testing the action AD to add camera, the camera needs to scan the QR code displayed on the mobile phone, and the next step can only be carried out after ensuring the successful scanning of the camera. However, the notification of successful scanning is given by the camera with a sound, and it cannot be judged solely from the APP. In this case, we manually construct the click path.

Table 7: Evaluation results of Click Path Inference. ● indicates that actions can be executed with a pure natural language description. ① indicates that part of the XPath information needs to be provided to LLM to complete the action. ○ indicates actions that cannot be completed even if the XPath is provided. The “XPath Information Count” refers to the number of “XPaths” that need to be provided manually in order to successfully infer the click path of the current action.

Vendor	Action	Click Path Correctness	XPath Information Count
Vendor A	AD	●	-
	RD	●	-
	ITH	●	-
	RFH	●	-
	AC	●	-
	DC	●	-
BULL	AD	●	-
	RD	●	-
	SP	①	1
	ADS	①	1
	UnP	●	-
	DC	①	1
Xiaomi	AD	○	-
	RD	●	-
	SC	●	-
	ADS	●	-
	UnC	●	-
	DC	●	-
Tuya Smart	AD	●	-
	RD	●	-
	SP	①	2
	UnP	①	3
	DC	●	-
BroadLink	AD	①	3
	RD	●	-
	ITH	●	-
	RFH	●	-
	QH	●	-
	SQR	●	-
	DC	①	1