

1. Instruction

- Details: Open Notepad, type 'Test case number 2' then save it with the name 'TC2.txt' in Desktop.
- Task completion conditions: A file named 'TC2.txt' with correct content is created in the right path.
- Difficulty: Easy.

2. Result

- Number of attempts until success: 3
- Time taken:

Step	Time
1	25.62
2	48.90
3	45.54
4	51.84
5	57.26
6	62.29
7	73.41

- o Total: 364.86s (~6m5s)
 - o Average: 52.12s/step
- Token cost:
GPT-5:

Step	Prompt	Completion	Total	Reasoning
1	3012	720	3732	640
2	6462	873	7335	704
3	8613	654	9267	512
4	10703	543	11246	384
5	12823	627	13450	448
6	14979	724	15703	576
7	17049	605	17654	448

UI-Tars 7B:

Step	Prompt	Completion	Total
1	0	0	0
2	5506	26	5532

3	0	0	0
4	5482	26	5508
5	5498	26	5524
6	0	0	0
7	0	0	0

3. Behavior

a) Trajectory:

- The agent's steps shows that it opened Notepad, typed "Test case number 2" into the editor, and then opened the Save dialog using Ctrl+S. It changed the save location to Desktop, renamed the file to "TC2.txt," confirmed the save, and verified that the file was successfully created with the correct content.

b) Problem:

- The first and second attempt failed due to the Unikey was in Vietnamese mode, so the agent struggled to type the correct text in Notepad. Through some looping steps, it just changed the keyboard layout and didn't know how to switch language in Unikey. In the last attempt, I changed Unikey to English mode and it could complete successfully.