Be the Agent

Play an AI agent with cards and find out what the model does, what the tools do and what carries the messages between them

An AI agent is a large language model (LLM) that can only write text: everything it does in the world happens through tools, which a program around it, the harness, runs on its behalf. The model knows only what is in its context, so its work can only be as good as the results that reach it.

Short description

Your group plays an AI agent with cards instead of a computer: one of you is the language model, one carries the messages, four of you are the tools, and one keeps the log. By the end you will have run by hand the loop that every AI agent runs on.

Steps

Basic experiment

1. Hand out the role cards; the Model's stack starts with the system card and Task 1 on top.
2. Play Task 1, then swap roles and play Task 2.
3. The Observer reads the log aloud: rounds, tool calls per tool, errors and cards on the stack.
4. Play Task 3 if there is time, then compare your log with the log of a real agent.

While you play, four rules apply:

- The Model sees only its stack of cards, and nobody talks to it.
- It writes one card at a time, starting with `THINK:`, `TOOL:` or `ANSWER:`.
- The Harness carries every `TOOL:` card to the right tool and brings back the `RESULT:` card.
- A tool does exactly what the card says, nothing more, and never guesses what was meant.

Extensions

Each extension changes the game in one respect, so that you can compare the result with your first round.

1. **Limited memory:** the stack holds at most eight cards, and a `SUMMARY:` card replaces the oldest ones when it is full.
2. **Two tools at once:** the Model may write two tool cards per turn, which the Harness carries in parallel.
3. **A new tool:** someone becomes `forecast("date")`, added to the system card, and you watch whether the Model uses it unprompted.
4. **A wrong result:** one tool secretly returns a wrong value, and you find out whether the Model notices.
5. **Be the tool for a chatbot:** solve Task 2 with a chatbot in your browser, supplying every file, page and calculation yourself.

Discussion

Write down your own answers first, then discuss them in your group.

1. In which round would the Model have been stuck without the tools, and which tool was used most?
2. What would have gone differently if the Model had had to give up its stack after ten cards?
3. Who did the thinking, who did the acting, and which of it could the Model actually do?

Concepts

- AI Agents: the loop of thinking, acting and observing that you just ran by hand.
- The Harness: the program around the model, which you played as the Harness.

Material

- A group of seven; with five or six, one person takes two roles.
- The cards, seven A4 sheets to cut along the dashed lines; print the sheet with blank cards several times, because a game uses twenty or more.
- The Observer's sheet, one per group and task.
- A pen for everyone, a calculator and room for the tools to sit apart from the Model.

All farm data and web pages are invented for this game and are not advice for real fertilisation.

Cards

Observer's sheet

One A4 sheet per group and task, with room for 22 rounds and the counts at the end.