Asteroid Form Challenge

Objective

Build a **browser agent** that reliably fills out a **mock insurance form.**

The forms are programmed to accept a specific set of demo data. Your agent should be given this data, and attempt to fill out the form in order to receive a code.

There are two response codes:

- ASTEROID_1: your agent successfully filled in the form with the correct information
- ASTEROID_0: your agent successfully filled in the form but there was a mistake in the data that it entered

You can use these forms for testing your agent. Try the easy form first, and then try

- **Hard Form**: https://asteroid.ai/form // Hard form demo data
- **Easy Form**: https://asteroid.ai/form2 // Easy form demo data

---

Challenge Details

- Your agent should parse the input and fill out the selected forms end-to-end.
- We will test your agent by calling it with a variety of input cases.
- There is **no predefined recipe**—you're free to explore different strategies (prompt engineering, tool use, chaining, evaluation loops, etc.).

Requirements

- You may use any framework or tools you find helpful, but minimal dependencies are preferred. We recommend starting from OpenAI's reference CUA implementation.
- Code should be **clear, modular, and readable**.
- Focus on **robustness and reliability**.
- Describe strategies taken to improve agents' performance.

Resources

- [Reference CUA OpenAI implementation](#)
- [OpenAI CUA Docs](#)
- [Anthropic Guide: Building Effective Agents](#)
- [Anthropic CUA Demo Code](#)
- Alternative frameworks (DOM-based, non-vision based):
    - https://github.com/browserbase/stagehand
    - https://pptr.dev/

Estimated Time Commitment

- **~8 hours** total

Submission & Logistics

- We'll provide an **OpenAI API key** with limited CUA credits (<$100).
- Push your code to a **private GitHub repo** and add:
    - `Joehewett`
    - `Mlcocdav`
- Feel free to reach out with any **questions or blockers**—we're here to help.
    - For example using local chromium browser might get stuck on dropdowns
- When complete, **send us the GitHub repo link**.