AI Close Encounters of the Third Kind 👽

Yair Friedman

Who am I

- ▶ Enjoying coding in C++ for three decades

- ▶ Software Security Professional

- ▶ Co-author and instructor of several Modern C++ trainings

- ▶ Member of Israel National body for C++ standard

yair.f.lists@gmail.com

Agenda

- "AI can write code for you" or "Vibe Coding"

- The first **ali**en

- The second **ali**en

- The third **ali**en

- Conclusions?

"AI can write code for you" or "Vibe Coding"

- ▶ Non-trivial but general design combination issue
- ▶ Simple search didn't provide any clear solution
- ▶ SOOP (Stack Overflow Oriented Programming) did not discover anything as well
- ▶ One blog with some complex, circular and partial explaination
- ▶ "AI can write code for you" or "Vibe Coding"
  - ‑ Implementing boiler-plated code
  - ‑ Implementing small, well-defined tasks
  - ‑ Creating lovley images

First **ALI**EN — ChatGPT

ChatGPT

👤 *Clear and detailed prompt of the code required*

💡 Goal... We want: *detailed accurate description*

✅ Step-by-Step: …

C++ code

⚠️ Notes: This is a minimal, educational implementation.

👤 *Indeed a very partial implementation*

👤 *Did not progress very far. Code is 30% functional*

Second **ALIE**N — Copilot

Copilot

Clear and detailed prompt of the code required

Here's the key: *repeat in few words*

C++ code

Code looks promising, but executing it appears as a very partial implementation

Running the given example… what is missing?

Ah, great catch! The issue lies in … Let's fix that with a small tweak

C++ code

Copilot

Same Result. can you highlight the changes? i do not see any difference except for the comments

You're absolutely right to be skeptical—because the version I gave you is functionally identical to the previous one.

Doh🤦 … What goes wrong?

Explanation … Working fix: …

C++ Code

This version … *same problem*

Copilot

✅ Working fix: … Here's a minimal and working version …

Below is a fully working example … C++ Code

The last version worked. *Some fudemental functionality is still missing*

C++ Code

This segfault on the 2nd call to …

The core of the crash … The Fix: … C++ Code

Still segfault same location

Copilot

- Diagnosing the persistent segfault *lists possible reasons …*

- C++ code. Can you post that minimal reproducible example and point out the crashing line?

- unmodified last example: … *output and GDB stack dump*

- Here's what's still happening under the hood … *More explanations* ***stack overflow link***

- The fix: … Partial C++ code

- Still segfault…

Copilot

- Diagnostic C++ code seeded with prints

- Output of last diagnostic version, did not segfault: *output*

- Next Steps for Root-Cause Analysis

- ***instructions how to build with ASan***

- *ASan output*

- Below is what's happening, … *More explanations*

- C++ Code

Copilot

- Still segfault

- What's happening … *More explanations*

- Partial C++ code

- *Doesn't segfault, functionality incomplete*

- *Nine versions of questions and C++ code did not make significant progress. Code is 75% functional*

Third **ALI**EN — Gemini
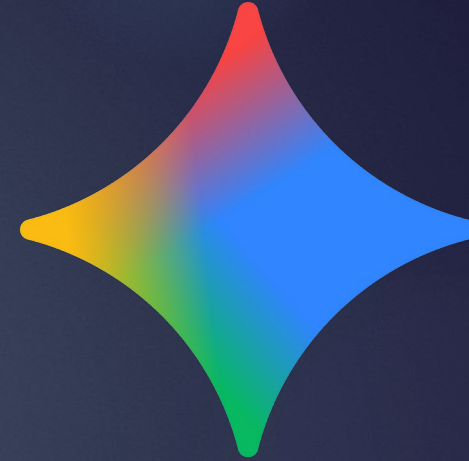
Gemini

Clear and detailed prompt of the code required

Very long detailed explanation of components... Implementation Strategy

C++ code

In a real application, you'd need... *list of missing parts*

Work slowly and meticulously. Please amend the above to meet these requirements: *list of missing parts*

The core challenge is... *pinpoint explanation of issue in hand:* **combination** ... C++ code

Verification of requirements *go over each point end explains how the code fulfills the requirement*

Gemini

The code doesn't compile: *error and possible fix*

More design questions

Thank you for your meticulous review. You've raised excellent points that highlight areas for improvement and clarification in the design.

Detailed design changes

C++ code *Well commented and added comments addressing design changes*

This code segfault

Output, GDB stack dump, ASan output

Gemini

- Thank you for providing the detailed bug report with both GDB and ASan output. This is extremely helpful. *Explanation on root cause …* The bug is **subtle** … *more detailed explanation* … The Fix:

- C++ code

- In update method … argument is corrupted

- You are absolutely correct. Thank you for catching this **critical** bug. *Explanation of corruption*

- C++ Code

Gemini

Code run as expected! 👏

Minor design questions and suggestions

Addresses the questions, but insists on its design. Code *is* 95% *functional*

Conclusions? And Notes

▶ When dealing with relatively simple tasks with many available solutions, all engines produces good output, when the information is scarce the picture is different. Stack Overflow (as well as reddit) is heavily used, but this is rarely admitted

▶ ChatGPT

→ Very partial implementation

▶ Copilot

→ Unstable code

→ Excellent diagnostic capabilities

▶ Gemini

→ Needs a very detailed prompt request

→ Frequent design changes

→ Delivered results

Conclusions? And Notes

- ▶ Getting code
  - → It might not compile
  - → If it does, it may still segfault
    - ⁻ Debugging capabilities for both diagnostic versions and memory analysis is very good
    - ⁻ Need to verify that the "fix" indeed fixes the issue in hand
  - → Need to verify that code is complete
- ▶ American-style interactions, less appealing to Israelis who want to get to the point

Q & A