

TASK 4[BONUS]

The LLMs failed significantly on most of the problems so I believe that very task solvable by an LLM is definitely solvable by a human.

There were many tasks which are easier for humans to solve but difficult for LLMs to solve, like almost all of the level 0 and 1 tasks are easier for humans to solve, and most of them can also solve level 2 tasks as well.

But LLMs were almost able to solve only half of level 0 tasks, and 2/5th of level 1 tasks, and almost no LLM was able to solve any level 2 task.

Here is a screenshot of the tasks I was able to solve on the SED PUZZLE website.

Get Started

Level 0 - Tutorial

Hello World

Hello World 2

Hello World 3

World Hello

Increment

Increment with Cursor

Decorate Me

Brackets

MISSISSIPPI

Level 1 - Easy

AB <-> BA

ABC ABC

ooo / xxx

AAAABBBB

Smiles

Growth

Peg Solitaire

Tower of Hanoi?

Tower of Hanoi? (without Poles)

Level 2 - Medium

123

19 Stars

Step Limit

Calculate in any order

Count Up

Amoeba

March

April

Binary Increment

Ternary Increment

Ternary Increment Impossible

Level 3 - Hard

Forbidden pattern

RPS Tournaments

Tower of JOIOI?

As you like

(Step Limit) As you like

As you like (Step Limit)