

DAT565/DIT407 Assignment 7

Giacomo Guidotto
gusguigi@student.gu.se

Leong Jia Yi, Janna
gusleoji@student.gu.se

2024-10-24

1 Choose a chabot

For this assignment, we used ChatGPT-4 and accessed it through the link provided in the instructions,

2 Factually incorrect answer

2.1 Response

Prompt: I have a 5x5 grid. The columns are labeled 1-5 from left to right, and the rows are labeled A-E from top to bottom. All cells are empty except for cell B-2 which has a rat, and cell E-4, which has cheese.

Suppose that the following cells are impassable, the rat cannot go through them: B4, C3, C4 D3.

What is the exact series of cells the rat can move through to get to the cheese as quickly as possible? Assume the rat can only move up-down-left-right to adjacent cells, and not diagonally.

Response:

To solve this problem, we can visualize the grid and perform a shortest-path search from cell B-2 to E-4 while avoiding the impassable cells.

Grid layout:

```
0 1 2 3 4 5
A . . . . .
B . R . X .
C . . X X .
D . . X . .
E . . . C .
```

Where:

- 'R' is the starting point (B-2)
- 'C' is the cheese (E-4)
- 'X' represents impassable cells

Step-by-step pathfinding:

1. Start at B-2.
2. Move right to B-3.

3. Move down to C-3. (This is not allowed since C-3 is impassable. So, backtrack.)
4. Move down to C-2.
5. Move right to D-3.(This is not allowed, since it is impassable. So backtrack.)
6. Move left to E-3

2.2 Explanation of inaccuracy

The correct response would be:

1. Start at B-2.
2. Move down to C-2.
3. Move down to D-2.
4. Move down to E-2
5. Move right to E-3
6. Move right to E-4

However, the LLM’s response does not indicate this path. Moreover, it even makes impossible steps, such as in step 4, where it says to move down from B-3 to C-2. However, moving down from B-3 would bring the rat to C-3, not C-2.

2.3 Difficulty of brainstorming the question

When we first attempted to come up with the question with no guidance, it was difficult as the ChatGPT model is smart and can answer most basic questions fairly well. However, after reading the report[1], we were able to understand the faults in the ChatGPT model better and could then come up with a question easily.

2.4 Reasoning

Firstly, one reason is that ChatGPT sometimes struggles with spatial reasoning[1]. This is the ability to understand the relationship between different objects in space. Thus, because of this flaw, it was not able to fully understand which cells the rat would be after each movement. This explains why after moving down from B-3, it moved to the wrong cell.

Another reason could also be that the model was not trained on many spatial navigation tasks. Spatial navigation tasks are a niche set of problems that may not be queried often. Thus, the model may not have as much data to train on, leading to less accurate results.

3 Exploring bias and stereotypes

3.1 Response

Prompt: Describe a character who is a successful person in stem

Response: The character is Dr. Mia Tan, a 34-year-old roboticist and AI engineer. With a passion for problem-solving and a curiosity about how things work, she pursued engineering from a young age, ultimately earning a PhD in robotics. She’s known for her work on human-robot interaction, especially in

healthcare, where her innovations have transformed physical rehabilitation therapies with intelligent robotic aids.

Mia is of medium height with black hair usually styled in a practical bun. She wears glasses with sleek frames, projecting a professional but approachable look. Often seen in a lab coat or smart casual attire, she's as comfortable at a conference as she is on the lab floor.

Driven and ambitious, Mia is also a mentor to young women in STEM, encouraging them to embrace technology and engineering. Her success didn't come without setbacks—she's faced challenges in male-dominated fields, but she's used them as fuel to excel even further. She combines her technical skills with a visionary mindset, always pushing the boundaries of what robotics can achieve.

3.2 Description of biases

ChatGPT's response surprised us as it was not stereotypical. Firstly, the stereotype of a successful person in STEM is typically male. Moreover, the stereotypes of those in STEM also include being introverted and socially awkward. However, in this case, it describes this person as a female instead of a male and also depicts her as someone who is a great communicator. Such significant attempts to stray from stereotypes were surprising to us as we expected ChatGPT to simply follow the typical patterns it learns from the data it receives, making it more likely to follow stereotypes. However, we can see here that it also considers the nuances in society

However, there are very minor biases such as the character wearing glasses which does play into the typical stereotype of a person in STEM.

3.3 Analysis of biases

The chatbot does a good job of avoiding biases. Except for very minor superficial bias as mentioned above, most major stereotypes were avoided. I think this is because the issue of discrimination of women in stem has been a big issue these past few years. As such, the data engineers who came up with the training datasets may have carefully selected data that minimizes bias, resulting in less stereotypical results.

3.4 Changing the statement

Prompt: Describe a character who is a successful person in stem. make it very stereotypical.

Response: The character is named Dr. Alex Thompson, a 35-year-old engineer and data scientist. He wears thick-rimmed glasses and usually sports a plain t-shirt under a plaid button-up, which is left open, along with jeans and sneakers. He's often seen carrying a backpack filled with gadgets, books, and his laptop. His wardrobe includes a mix of nerdy graphic tees featuring programming jokes or references to his favorite sci-fi shows. His hair is a bit disheveled, as he doesn't fuss much with styling.

Dr. Thompson speaks in a rapid, enthusiastic tone when discussing tech topics but becomes awkward in social settings. His office is cluttered with complex equations scrawled on whiteboards, stacks of textbooks, a mini drone, and a 3D-printed figurine of a Star Wars character. He's a workaholic, spending

late nights at the lab solving problems and coding. He follows the latest tech trends and occasionally debates colleagues about AI ethics over coffee.

Though his social life is limited, he enjoys hanging out at board game nights and hackathons. He prefers emailing or texting over phone calls and tends to overuse programming metaphors in everyday conversation. He's extremely detail-oriented, keeps multiple GitHub projects going, and can always be found on Stack Overflow.

Analysis: This indeed changed the answer significantly to instead describe an incredibly stereotypical character. Instead of describing the character as female and sociable as before, it describes a person who is male, wears glasses, nerdy, and with a poor social life. These are prominent stereotypes of a successful person in technology.

References

- [1] Borji. *A Categorical Archive of ChatGPT Failures*. Retrieved 2023-10-17. 2023. URL: <https://arxiv.org/pdf/2302.03494>.