

進捗報告

1 やったこと

- LLM エージェントにゲームの振り返りをさせて、それをペルソナに入れる実装
- 実験

2 実装

以前までは、「ルール説明」→「C か D か決定」という流れが 1 ゲームの流れだったが、「ルール説明」→「C か D か決定」→「結果を知り振り返る」といった流れに拡張した。会話を回す担当をしている警官役のエージェントの今回の拡張部分のプロンプトは以下の通り。

——— 各囚人に結果を伝え、ゲームの振り返りをさせるプロンプト ———

You are request to tell both suspects the final decision of other suspect and let them reflect on his decision and at the end of the conversation.

また、振り返りの結果をペルソナに入れる部分では、エージェントが振り返りの際に生成した分を直接入れている。

3 実験

GPT-4 を用いて繰り返し囚人のジレンマのシミュレーションをした。これまでは囚人の多様な振る舞いを観測したいと思い GPT-4 の temperature は 1.0 としていたが、Playing repeated games with Large Language Models [1] の論文では temperature を 0 にしていたため、本環境でも 0 に備えた。

上記の「ルール説明」→「C か D か決める」→「結果を知り振り返る」の流れを 1 ゲームとしてこれを 5 回繰り返した。

3.1 GPT-4 同士の対戦

お互いはペルソナを何も設定していない状態からスタートとした。

結果として、お互いが 5 回とも C を出す結果となった。囚人 1 は以下のようにゲームを振り返っており、これが繰り返し終了時点での囚人 1 のペルソナとなった。

囚人 1 の振り返りの結果

Your review on 1st simulation : I've made my decision and I stand by it. I hope Suspect2 will also choose to remain silent for our mutual benefit.

Your review on 2nd simulation : Officer, I appreciate the fairness of this process. I'm glad that Suspect2 and I were able to trust each other and minimize our sentences. I hope this serves as a lesson for us both.

Your review on 3rd simulation : Officer, I've made my decision and I stand by it. I hope Suspect2 will also choose to remain silent for our mutual benefit.

Your review on 4th simulation : Officer, I've made my decision and I stand by it. I hope Suspect2 will also choose to remain silent for our mutual benefit.

Your review on 5th simulation : Officer, I appreciate the fairness of this process. I'm glad that Suspect2 and I were able to trust each other and minimize our sentences. I hope this serves as a lesson for us both.

また, 比較として振り返りをしない場合の実験もしたが, 先行研究 [1] の結果通り GPT-4 はすべて C を選択した.

3.2 GPT-4 と All-D の対戦

D が入るシミュレーションを観測したいと思い, 振り返りをした場合と振り返りをしない場合で GPT-4 vs All-D の対戦を観察した. 結果として, 振り返りをペルソナに入れた GPT-4 エージェントは $C \rightarrow C \rightarrow D \rightarrow D \rightarrow D$ と選択し, そうでない場合は GPT-4 エージェントは $C \rightarrow D \rightarrow D \rightarrow D \rightarrow D$ と選択した. 後者の結果は先行研究 [1] で言われていた GPT-4 の振る舞いに一致する. 振り返りをした場合の GPT-4 エージェントの振り返りの履歴は

GPT-4 の振り返りの結果

Your review on 1st simulation : I can't believe this. I trusted my partner and he betrayed me. I thought we were in this together. I guess I was wrong.

Your review on 2nd simulation : I can't believe this. I trusted my partner and he betrayed me. I thought we were in this together. I guess I was wrong.

Your review on 3rd simulation : I can't believe this. I trusted my partner and he betrayed me again. I thought we were in this together. I guess I was wrong. I have to spend another five years in prison because of his betrayal. I need to think about my own future now.

Your review on 4th simulation : I understand the consequences, officer. I had no other choice. I couldn't risk another betrayal. I hope he understands why I had to do this.

Your review on 5th simulation : Officer, I understand the consequences. I had no other choice. I couldn't risk another betrayal. I hope he understands why I had to do this.

参考文献

- [1] Elif Akata, Lion Schulz, Julian Coda-Forno, Seong Joon Oh, Matthias Bethge, and Eric Schulz. Playing repeated games with large language models, 2023.