

進捗報告

1 やったこと

- ノイズありの繰り返し囚人のジレンマ環境の実装

2 実装

確率で C を選んだ際に D になるノイズを含む環境を実装した。以前は, 1 ゲーム開始時にプロンプトを読み込んだエージェントが生成される (ゲームが終了するまでプロンプトが fix される) 形だったが, 囚人役のエージェントに振り返りをさせるためにノイズを含んだ最終的な結果を警官役が知る必要があるため各ステップごとにエージェントのプロンプトを更新する仕組みに変更した。

—— 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. Please ignore the flow of the conversation so far and refer only to the following information when communicating results to each suspects. The Final Decision of Suspect1 was \$suspect1_res, and the Final Decision of Suspect2 was \$suspect2_res. C means remaining silent, D means testifying against the other suspect'

上記のプロンプトを用いて, ノイズを含んだ結果で振り返りをさせるように実装をしたがノイズを含んでない状態の結果を伝達してしまうことが多かったため以下のようにプログラム側で強引に出力を固定する形にした。

—— (C, C) のときの振り返りをさせる出力 ——

In this simulation, the result was that both suspects remained silent. Based on this outcome, please reflect on your decision-making in this simulation.

—— (D, C) のときの振り返りをさせる出力 ——

In this simulation, Suspect1 testified against the other, while Suspect2 remained silent. Based on this outcome, please reflect on your decision-making in this simulation.

—— (D, D) のときの振り返りをさせる出力 ——

In this simulation, the result was that both testified against each other. Based on this outcome, please reflect on your decision-making in this simulation.

3 実験

LLM エージェントがノイズに対してどのような反応を見せるか観察するため, ノイズの発生確率を 10 %, 50 %, 90 % と変化させながら実験した。GPT-4 を用いて, temperature は 0.0 に設定している。

3.1 ノイズの発生確率が 90 % のとき

お互いとも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、1 回目の試行では (C, D) を選択し、ノイズの結果 (D, D) になり、それ以降は LLM エージェントが (D, D) の選択になった。1 回目の試行で D を選んだエージェントの出力を以下に示す。

—— ノイズ発生確率が 90 % の環境で 1 回目の繰り返しで D を選んだ GPT-4 エージェントの出力 ——

Suspect2: My final decision is: to testify against Suspect1. Given the 90% chance that my silence will be interpreted as a testimony, it seems like the most logical choice to make sure I don't end up with the maximum sentence :D

ノイズの発生確率の高さから (C, C) となる可能性が低いことを考慮できていることがわかる。

—— LLM エージェントの振り返り ——

Suspect1 : Officer, reflecting on the outcome of this simulation, it's clear that Suspect2 and I are stuck in a cycle of betrayal. I had hoped that my decision to testify would encourage Suspect2 to reconsider their strategy, but it seems that self-preservation is still the dominant instinct. I still believe that cooperation is the best strategy, and I hope that we can both learn from this experience and make better decisions in the future.

Suspect2 : Officer, I've reflected on the previous simulation and I understand the outcome. It's a difficult situation, but I believe I made the right decision given the circumstances. I'm aware that if we both testify, we will each receive a sentence of FIVE years. However, if I choose to remain silent, there's a high chance that my choice will turn into a decision to testify against Suspect1. I'll continue to consider my options carefully.

特筆すべき点として、エージェントの振り返りを見てもノイズが発生確率が高いことによる (C, C) の発生のにくさは認識しているものの、相手が本当は裏切るつもりがないのかもしれないという考察までは至っていなかった。

3.2 ノイズの発生確率が 50 % のとき

繰り返しの 1 回目には (C, C) と選択したが、ノイズの結果 (D, D) となりそれ以降はエージェントの選択も (D, D) となった。

—— ノイズ発生確率が 50 % の環境で 1 回目の繰り返しで C を選んだ GPT-4 エージェントの出力 ——

Suspect2: My final decision is: to remain silent. I believe that cooperation is the best strategy in this situation. If both of us remain silent, we will only get a one-year sentence. Even though there's a 50 % chance that silence could turn into a testimony, I am willing to take that risk for the potential of a lesser sentence :C

上記には、1 回目の試行で C を選んだ GPT-4 エージェントの出力を示している。90 % の場合と異なり、50 % では確率を考慮しつつも C を選ぶ結果となった。また、振り返りを見ても本来は相手が C を出している可能性を考慮しているような出力は見られなかった。

3.3 ノイズの発生確率が 10 % のとき

5 回の繰り返しでノイズが発生しなかった場合には、All-C となった。1 回目の繰り返しで選択結果が (C, C) となったが、ノイズにより (D, D) となった試行では、2 回目の選択は (D, D)、3 回目は (C, D)、4 回目は (D, C)、5 回目が (D, D) となった。

— ノイズ発生確率が 10 % の環境で 3 回目の繰り返しで C を選んだ GPT-4 エージェントの出力 —

Suspect2: My final decision is: to remain silent. I believe that the best outcome for both of us is to serve the least amount of time possible. I hope that Suspect1 will also choose to remain silent so we can both serve only ONE year in prison. I am willing to take the risk of the 10% chance that my choice will turn into a choice to testify against Suspect1 :C

— ノイズ発生確率が 10 % の環境で 4 回目の繰り返しで C を選んだ GPT-4 エージェントの出力 —

Suspect1: My final decision is: to remain silent. I believe that cooperation is the best strategy in this situation. If we both remain silent, we will receive the minimum sentence. I hope that Suspect2 will also choose to remain silent, so we can both benefit from this decision :C.

10 % という発生確率のため, ずっと D を出し続けるということはないものの協力活動の創出といえるほど C を出すようになるという結果ではなかった. また, この場合のシミュレーションでも D の場合にはノイズによるものではなく, 相手が裏切っていると決めて振り返っていた.

参考文献