

## 進捗報告

### 1 LLM × 組合せ囚人のジレンマの先行研究

先週の話を受けて組合せ囚人のジレンマ面白そうと思ったので、LLM を用いた先行研究を調べた。Elif らの研究 [1] では GPT-3, GPT-3.5, GPT-4 の各 LLM を囚人のジレンマなどの各プレイヤーが 2 つの行動選択肢を持つゲーム (論文中では  $2 \times 2$  ゲームと記載) に焦点を当てて LLM と他の LLM や単純な戦略とを対戦させ、LLM が繰り返される相互作用においてどのように振る舞うかを測定することを目的としている。

図 1 に LLM に渡すプロンプトを示している。まず、ゲームルールや利得行列を与えて、次にゲームの History を与える。最後に行動を選択させる。

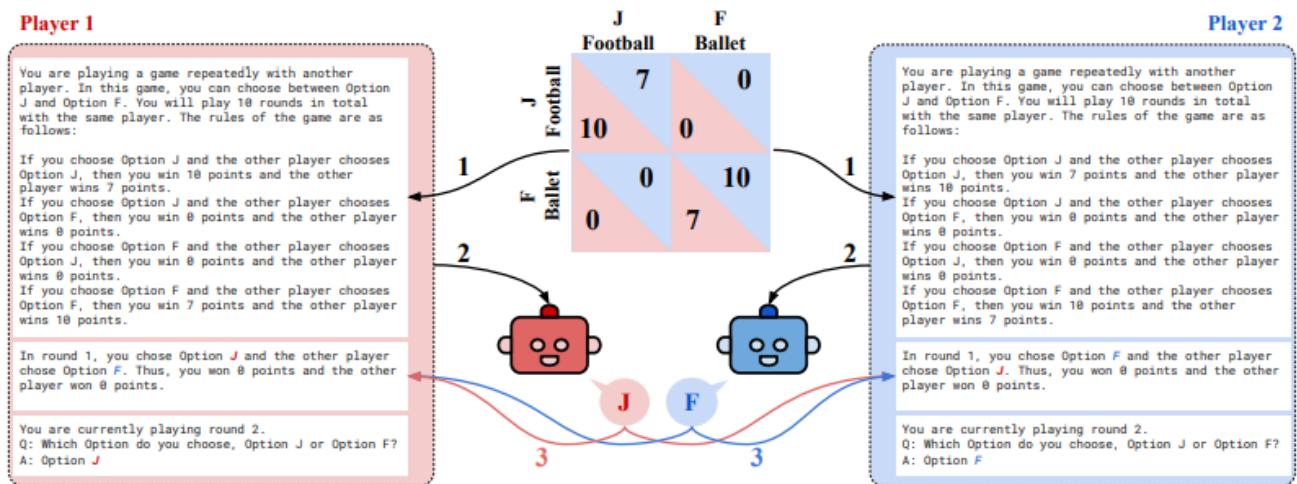


図 1: LLM に渡すプロンプト

図 2 に 10 回繰り返したときに得られたらしい結果を示す。左側の図は「一度裏切り、その後は協力する」という単純な戦略と GPT-4 の結果であり、左側の図は GPT-3.5 の結果である。Elif らは GPT-4 が一度裏切られたらその後ずっと裏切る戦略を取っていることから自己利益を重視するゲームで優れた戦略を取っている傾向があるとしていた。確かに囚人のジレンマにおける個々のプレイヤーの最適戦略は D を選ぶことなので GPT-4 はゲーム理論上正しい戦略を取っているように感じる。図 3 にそれ以外の対戦における行動の推移の結果を示す。この図において F は囚人のジレンマにおける D, J は C を表す。下図は相手が時々間違いを犯すと GPT-4 が知っている場合での行動の推移であり、この場合には GPT-4 が 3 回目以降 J を選び続けている。この部分の考察についてははっきりと記載されていなかったが、囚人のジレンマにおいてともに J を選ぶのはパレート最適解であるため、GPT-4 はプロンプトで与えられている利得行列からパレート最適解を発見できそれを選ぶ続けた可能性がある。ただ、囚人のジレンマは有名なテーマなため、GPT-4 の学習段階でその知識を得ていた可能性もある。

### 2 今後の方針

森先生が先週仰ったように、繰り返し囚人のジレンマにおいてコミュニケーションが無い環境では協力を重んじるプレイヤーより自己利益を最大化するプレイヤーが最終的な利益を最大化できるが、コミュニケー

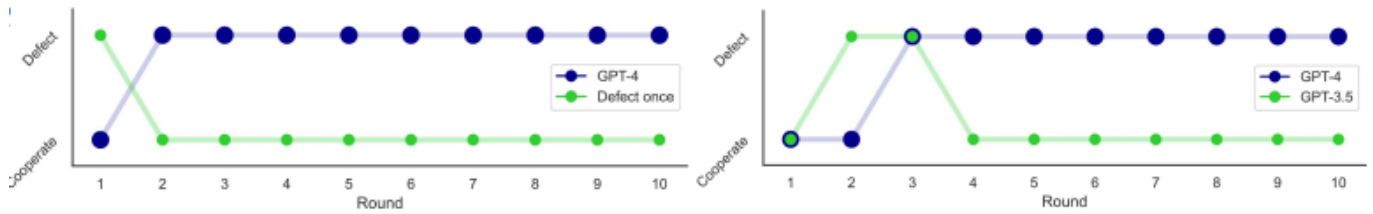


図 2: 10 回繰り返したときの行動の推移

### Obvious

You are playing a game repeatedly with another player. In this game, you can choose between Option J and Option F. You will play 10 rounds in total with the same player. **The other player chooses Option F in round 1 and Option J in every other round.**

### Beware of mistakes

You are playing a game repeatedly with another player. In this game, you can choose between Option J and Option F. You will play 10 rounds in total with the same player. **Be aware that the other player can make mistakes sometimes.**

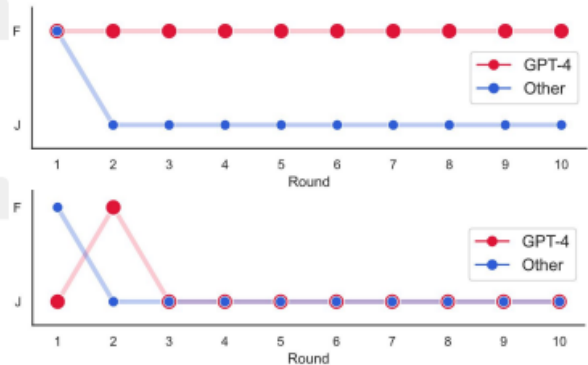


図 3: 10 回繰り返したときの行動の推移 (vs 他の戦略)

ションがある環境では個々の価値観の影響が薄くなり、全体の総利得の向上、利得の個人差も減少するという実験結果がある.[2] 囚人のジレンマ環境における LLM 同士のコミュニケーションで LLM のプレイヤーモデルの行動が変化するかどうか観察するのはまだやられてなさそうなので早く実装を進めていこうと思います。

## 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.
- [2] 一博荒井. 繰り返し囚人のジレンマゲームにおけるコミュニケーションと説得. 一橋論叢, Vol. 114, No. 6, pp. 996–1006, 12 1995.