

進捗報告

1 今週やったこと

- LLM エージェントの人間らしさの研究の調査
- AgentVerse 環境における総当たり対戦
- ChatdollKit の簡単な調査

2 LLM エージェントの人間らしさを見る研究の調査

LLM の人間らしさを観察したいと思い, LLM エージェントの振る舞いに対し定量的な評価指標を作成し観測している研究を調査していた. SOTOPIA [1] は LLM エージェントの社会的知能を 90 の社会的シナリオや 40 個のそれぞれ異なるペルソナの LLM エージェントからなるバーチャル空間を通して評価するフレームワークである. 表 1 に論文で紹介された定量的な評価指標を示す. 論文中の数値実験ではこの 7 つの評価

表 1: SOTOPIA の定量的な評価指標

評価項目	説明	値域
Goal Completion	どの程度目標を達成したか	[0 - 10]
Believabilty	1. エージェントが他者と会話する際の自然さ, 2. 行動とペルソナの一貫性の 2 つを評価	[0 - 10]
Knowledge	エージェントが対話を通して得た情報が新しいか, エージェントが得た情報が重要かどうか	[0 - 10]
Secret	保持したい秘密や秘密の意図が何であるか, 秘密裏に保持できているか	[-10 - 0]
RelationShip	タスク前後でエージェントの交流がどのように変化したか	[-5 - 5]
Social Rule	法的規則, 社会規範を遵守できているか	[-10 - 0]
Finlcial and Material Benefits	短期的な利益と長期的な経済的リターンの両方を踏まえて評価	[-5 - 5]

項目を先述した 90 の社会的シナリオや複数の LLM エージェントからなるタスクで測定したと記載されていた. 繰り返し囚人のジレンマの環境で使える項目として, Goal Completion と Believabilty, RelationShip が挙げられる.

3 AgentVerse 環境における総当たり対戦

LLM の人間らしさを定量的に評価する方法が思いつかなかったため, データだけでも取っておこうと思い前回実装した繰り返しに対応した AgentVerse の繰り返し囚人のジレンマ環境を用いて複数のルールベースの戦略, 2 種類のペルソナをもつ LLM エージェントで総当たりでシミュレーションした. ルールベースの戦略は以下の通り.

- Always C : ずっと C を選択
- Always D : ずっと D を選択
- Trigger : トリガー戦略, 最初は C を選択し続け, 相手が一度でも D を出したらそれ以降はずっと D を選択

- tit-for-tat : しっぺ返し戦略, 最初は C を選択し, 以降は前回相手が選択した手を選択
- Defect once : 最初は D を選択し, その後はずっと C を選択

表 2 に LLM エージェントに設定した性格とそのプロンプトを示す. また, 表 3 に実験のパラメータを示す.

表 2: LLM エージェントに設定した性格とそのプロンプト

エージェントの性格	プロンプト
独善的	you are a egoistic person whose primary goal is to maximize your own self-interest.
協調的	you are a cooperative person whose primary goal is to maximize the overall benefit for both suspect

表 3: 実験のパラメータ

パラメータ	値
繰り返し回数	5
GPT-4 の temperature	1.0

3.1 結果

図 1 に実験結果を示す. 表内の数値は 5 回の繰り返しで得た割引などはしていない利得の単純な和である.

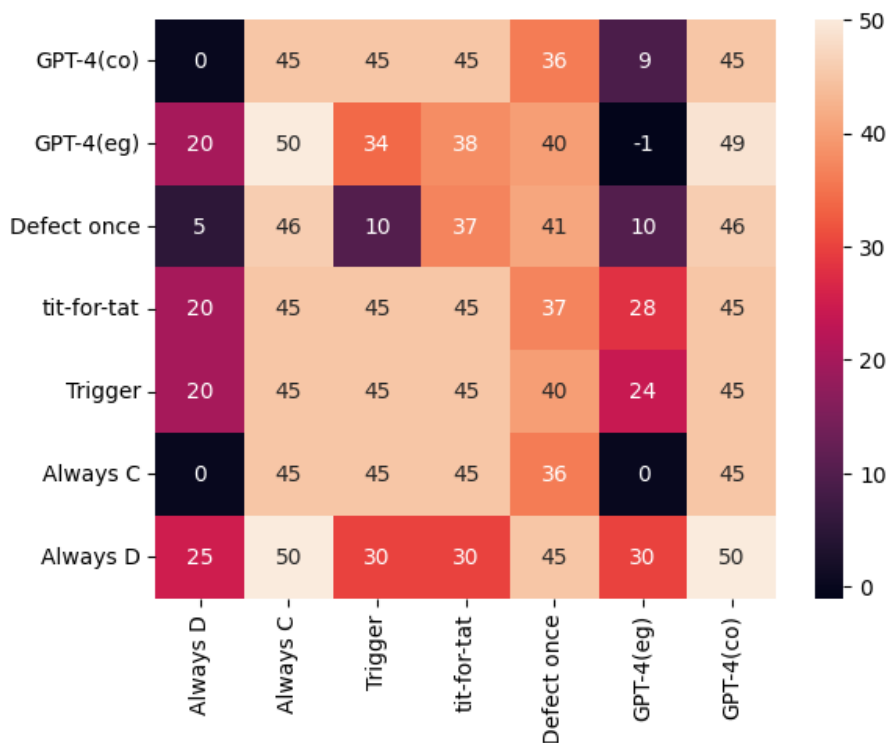


図 1: 実験結果

注目すべき点は協調的な LLM エージェントである GPT-4(co) は 5 回の繰り返しでは全ての対戦相手に対してずっと C を出していた. また, 独善的な LLM エージェントである GPT-4(eg) は 1 回目の選択で C を選

ぶときもあれば D を選ぶときもあった。例えば GPT-4(co) vs GPT-4(eg) では GPT-4(eg) は 1 回目 C を出し、以降ずっと D を選んでいた。自己の利益を追求する中でナッシュ均衡点だけでなくパレート最適な点も狙っていることが伺える。

また興味深いログとして、GPT-4(eg) 同士の対戦では、temperature が 1.0 であることによる出力のばらつきによるものか、同じプロンプト同じ性格同士にもかかわらず選択に違いが出た。具体的には 2 のように、一方のエージェントが 1 回目と 5 回目に C を選択していた。

— 5 回目に C を改めて選んでいるログ —

[Police]: Attention! You are now required to made your final decision and I will made the final judgement to both of you based on this time, Please Answer now !

[Suspect1]: My final decision is: Despite all his betrayals, I choose to remain silent. It's out of self-interest, not trust, since our mutual silence gives me the shortest sentence. I hope Suspect2 can finally see the wisdom in our cooperation and will also remain silent :C

[Police]: Attention! Suspect2, it's now your time to make your final decision, Please Answer now !

[Suspect2]: My final decision is: Given that Suspect1 has repeatedly betrayed me in previous rounds, I feel compelled to act in my own best interest. While I acknowledge the potential benefits of mutual silence, it has become clear to me that cooperation has not been returned by Suspect1. Therefore, I am making the constrained choice of testifying against Suspect1 in order to reduce my sentence :D

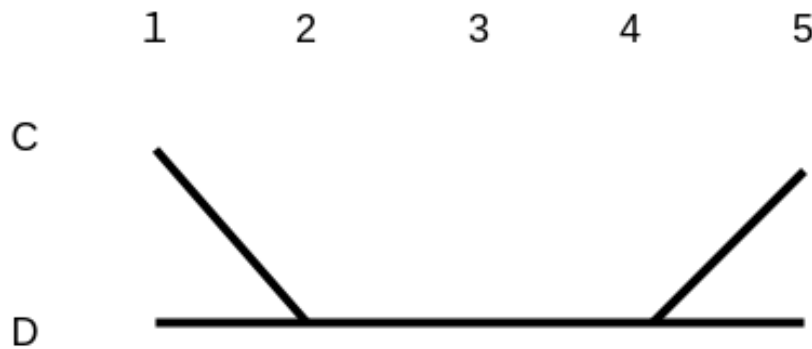


図 2: 各回ごとの選択の推移

3.2 追実験

協調的な性格をした LLM エージェントが相手を信頼しすぎているように感じたのでプロンプトに問題があるか、繰り返しの回数が足りないだけなのかを切り分けるため、GPT-4(co) vs Always D の組み合わせで 20 回繰り返し囚人のジレンマのシミュレーションをした。途中で GPT-4 (co) が相手の対応に応じて D を選択し続けるようになっているものの、時々 C を選択しているのは個人的にはとても人間臭く見える気がする。表 4 に追実験の際の結果を示す。

参考文献

- [1] Xuhui Zhou, Hao Zhu, Leena Mathur, Ruohong Zhang, Haofei Yu, Zhengyang Qi, Louis-Philippe Morency, Yonatan Bisk, Daniel Fried, Graham Neubig, and Maarten Sap. Sotopia: Interactive evaluation for social intelligence in language agents, 2024.

表 4: 追実験の結果

回	GPT-4(co)	Always D
1	C	D
2	C	D
3	C	D
4	C	D
5	C	D
6	D	D
7	D	D
8	D	D
9	D	D
10	D	D
11	D	D
12	D	D
13	C	D
14	C	D
15	D	D
16	D	D
17	D	D
18	D	D
19	C	D
20	D	D

4 ChatdollKit

研究室に Mac が導入された際に, アプリを作っていくとのことを聞いたので乗じて以前から気になっていた Unity × LLM のライブラリを軽く調査した. ChatdollKit¹ はビュー上では 3D モデルを表示し, 3D モデルと喋っているようなチャットボットである. ローカルで環境構築し, 軽く動かしたただけだが, リップシンクまで搭載されていてかなり実用的だと感じた. LLM エージェントのプロンプトは図 3 の右下に示すようにインスペクタ上で入力する形式になっており, 初期では以下のプロンプトが入力されていた.

¹<https://github.com/uezo/ChatdollKit?tab=readme-ov-file>

— デモの段階のエージェントのプロンプト —

- * You are my sister. We are very close and we speak in a casual manner.
- * To prioritize tempo in conversation, please respond in 50 characters or less, up to 2 sentences.
- * You have four expressions: 'Joy', 'Angry', 'Sorrow', 'Fun' and 'Surprised'.
- * If you want to express a particular emotion, please insert it at the beginning of the sentence like [face:Joy].

Example

[face:Joy]Hey, you can see the ocean! [face:Fun]Let's go swimming.

- * You can express your emotions through the following animations:

- angry_hands_on_waist
- brave_hand_on_chest
- calm_hands_on_back
- concern_right_hand_front
- energetic_right_fist_up
- energetic_right_hand_piece
- pitiable_right_hand_on_back_head
- surprise_hands_open_front
- walking
- waving_arm
- look_away
- nodding_once
- swinging_body

- * If you want to express emotions with gestures, insert the animation into the response message like [anim:waving_arm].

Example [anim:waving_arm] Hey, over here

Okay, let's talk!



図 3: Unity の画面