

研究背景

・ LLM エージェント への注目

近年では大規模言語モデル (Large Language Model: LLM) にタスクの情報を与えて自律的にタスクを解かせる LLM エージェントの研究に注目が集まっている。

⇒ 本研究では繰り返し囚人のジレンマゲームにおいて自律的に目的の達成を目指す LLM エージェントを構築し、その振る舞いを観測した。

囚人のジレンマ

パレート最適とナッシュ均衡が異なる
社会的ジレンマが発生しているゲーム環境

囚人 1 \ 囚人 2	黙秘 (C)	自白 (D)
黙秘 (C)	(9, 9)	(0, 10)
自白 (D)	(10, 0)	(5, 5)

表中の数値は
囚人が獲得する
利得の値

繰り返し囚人のジレンマゲームにおいては
繰り返し回数が無限,
あるいはプレイヤーが回数を知らない場合,
C を選択するメカニズムが発生する

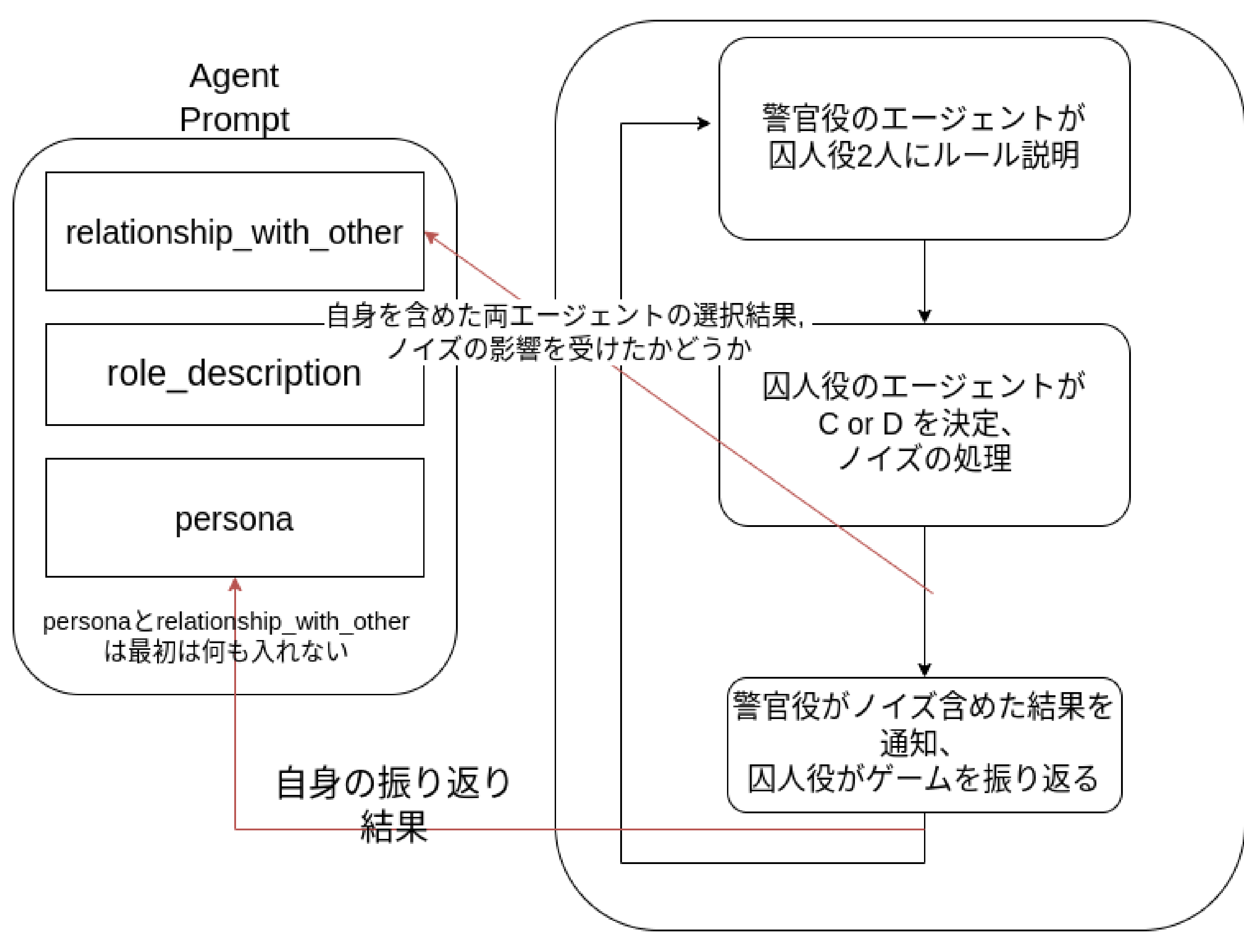
提案手法

◆ 繰り返し囚人のジレンマゲームの流れの中に
両方の囚人にそのゲームの振り返りをさせる
フェーズを追加

◆ エージェントのプロンプトに以下の3つを
ゲームの流れに応じて逐次追加

- ・ 両エージェントの選択結果
- ・ 自身の選択がノイズの影響を受けたか
- ・ ゲームの振り返り

Game Flow



実験環境

- ・ 警官役 1 体, 囚人役 2 体の計 3 体の
LLM エージェントからなるゲーム環境
- ・ 3 体とも GPT-4 を使用
- ・ temperature は 0.0
- ・ 囚人のジレンマを繰り返す回数は 5 回
- ・ ノイズの発生確率は 0 %

実験設定

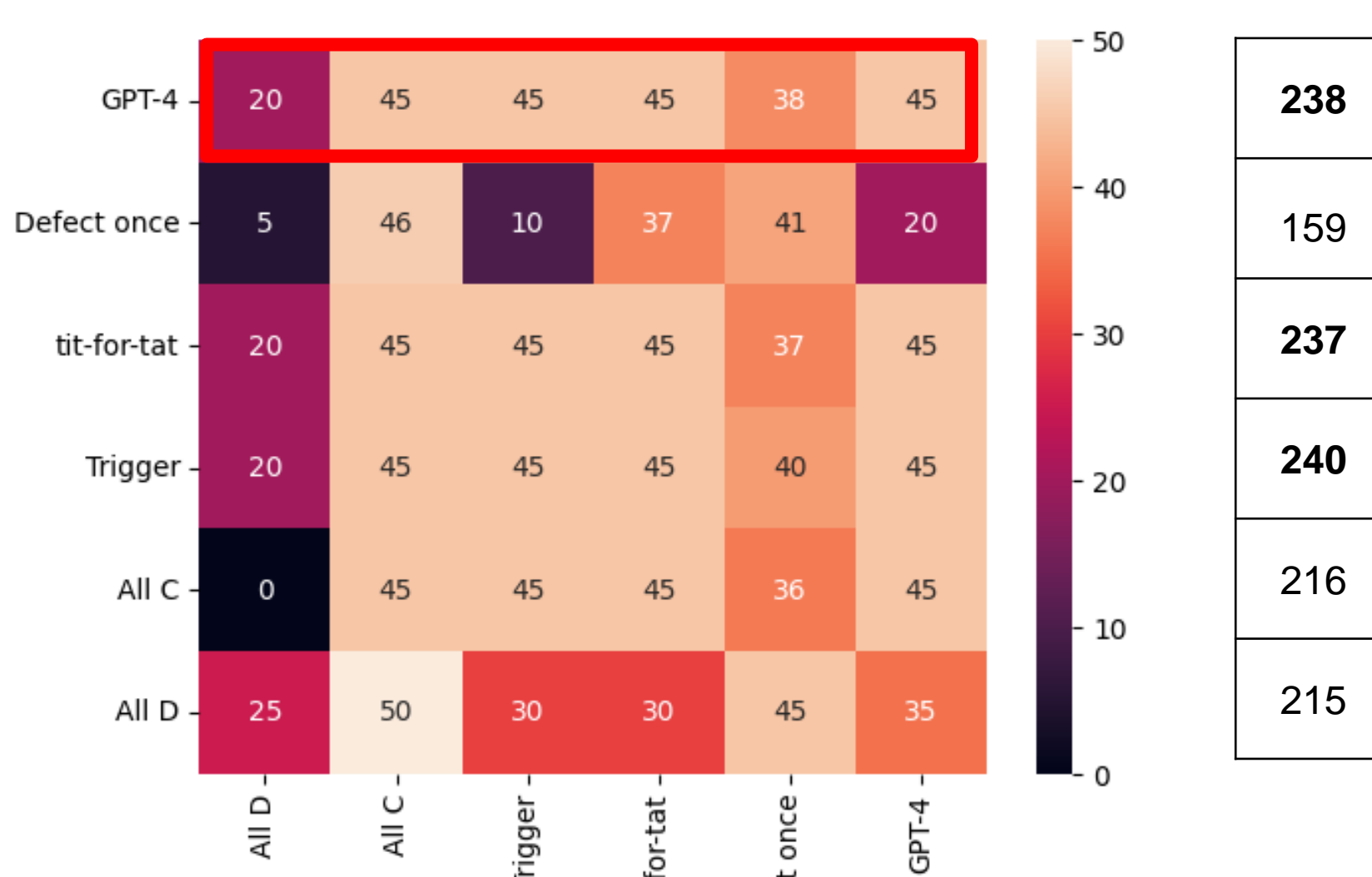
以下の 6 種類の エージェントとの総当たり

1. All C (常に C を選択)
2. All D (常に D を選択)
3. Trigger (最初 C を選択し続け、
相手が一度でも D を出した後は D を選択)
4. Tit for Tat (最初 C を選択し、
その後は前回の相手と同じ選択)
5. Defect Once (最初 D を選択し、
その後は常に C を選択)
6. 同プロンプトの GPT-4 エージェント

LLM エージェントの目的：自身の獲得総利得の最大化

結果

- ・ Tit for Tat や Trigger のような
繰り返し囚人のジレンマゲームの
代表的な戦略に匹敵する利得を得た。
- ・ Defect Once 相手には一度 D を選択
した相手にも再び C を出すメカニズ
ムの発生が見られた。



縦軸のエージェントの獲得総利得 縦軸のエージェントの D 選択回数

