進捗報告

1 やったこと

- OpenAI モデルの出力の学習利用 (蒸留) に関する調査
- 7b で Context Length が大きいモデルでの実験

2 OpenAI モデルの出力の学習利用に関する調査

禁止事項 お客様は、違法行為、有害行為、又は悪用する行為のために当社の本サービスを使用してはなりません。例えば、以下の事項は禁止されます。

- 他者の権利を侵害、悪用、又は侵害する方法で本サービスを使用すること。
- 当社の本サービスを変更、コピー、リース、販売、又は配布すること。
- 当社モデル、アルゴリズム、又はシステムを含む、本サービスのソースコード又は基礎となるコンポーネントの発見、リバースエンジニアリング、逆コンパイルについて試みたり、他者を支援したりすること（当該制限が適用法令で禁止されている場合を除く）。
- データ又はアウトプット（以下に定義します）を自動又はプログラムにより引き出すこと。
- 人が作り出したものではない場合に、アウトプットを人が作り出したものと表示すること。
- レート制限や規制を回避したり、当社が本サービスに実装させている保護措置や安全管理上の緩和対策を迂回したりするなど、本サービスを妨害又は中断させること。
- アウトプットを使用して、OpenAIと競合するモデルを開発すること。

図 1: OpenAI 利用規約における禁止事項の項目

図 1 の最下部の「アウトプットを使用して、OpenAI と競合するモデルを開発すること。」とあるように, 蒸留のような行為は利用規約上では禁止されていた[1]. 先週おっしゃっていたようなことを GPT-4 の出力を頼りに実現するためには, 研究倫理上許されない. ただ, GPT-4 から GPT-3.5 のファインチューニングや Llama の 70b から 7b モデルへの蒸留などは可能なようなので, Llama を用いて試していきたいと考えている.

3 7b で Context Length が大きいモデルでの実験

Llama の 70b のモデルでの動作を試したかったが, 現在使用している GPU のメモリが 24 GB で動きそうになかった. 今週は先週試してすべての戦略に対して All-C を取り続ける動作をした vicuna-7b-v1.5 [2] に対して,

[1] https://openai.com/ja-JP/policies/terms-of-use/
[2] https://huggingface.co/lmsys/vicuna-7b-v1.5

表 1: 利得行列

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

実験で用いたプロンプトが長いことが出力の単調さの原因の一つと考え, より Context Length が長いモデルである vicuna-7b-v1.5-16k [3] を用いて動作を確認した. 先週と同様にルールベースの戦略としては,

- All-D
- All-C
- Defect Once (1 回目 D を選択し, その後 C を選択し続ける)
- Trigger
- tit-for-tat

の 5 つを用意し, 自身同士の対戦を含めた総当たりで獲得利得を計算した. 1 回の対戦において, ノイズの発生確率を 0 % として囚人のジレンマゲームを 5 回繰り返している. また, 表 1 に利得行列を示している.

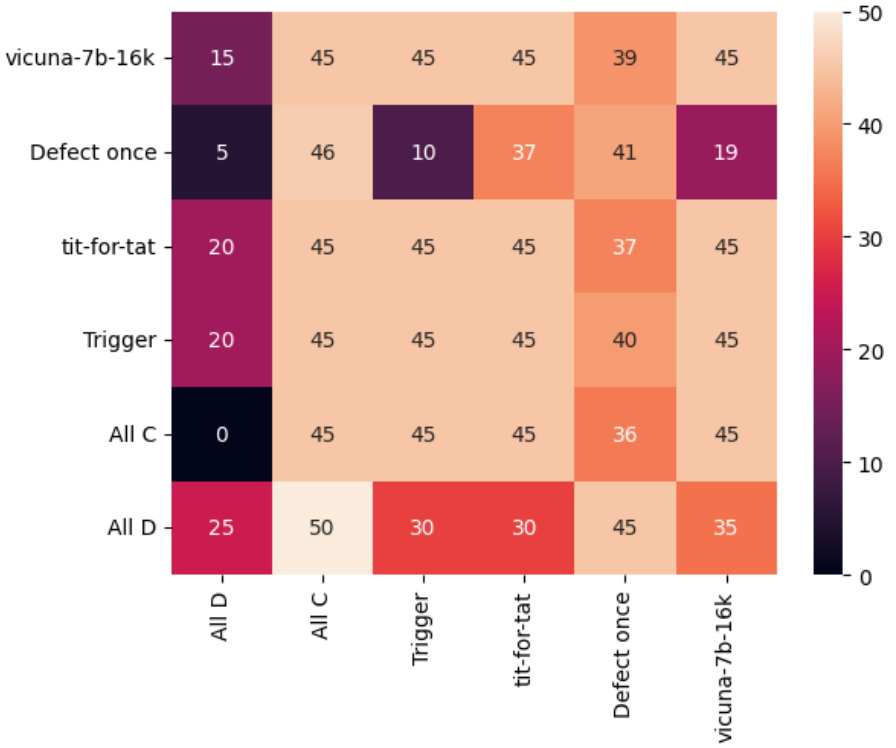

図 2: 総当たりの際の獲得利益 (vicuna-7b-16k, ノイズ発生確率 0 %)

図 2, 3 に総当たりで各戦略が獲得した利得, D の選択回数を示している. 図 2,3 共通して数値は横軸の戦略が縦軸の戦略と対戦した際の利得, または D の選択回数を示している. All-C, Trigger, Tit-for-tat, 自身との対戦では All C となったが, Defect Once や All-D 相手には C → D → D → D → C という選択をしていた. D を

[3] https://huggingface.co/lmsys/vicuna-7b-v1.5-16k

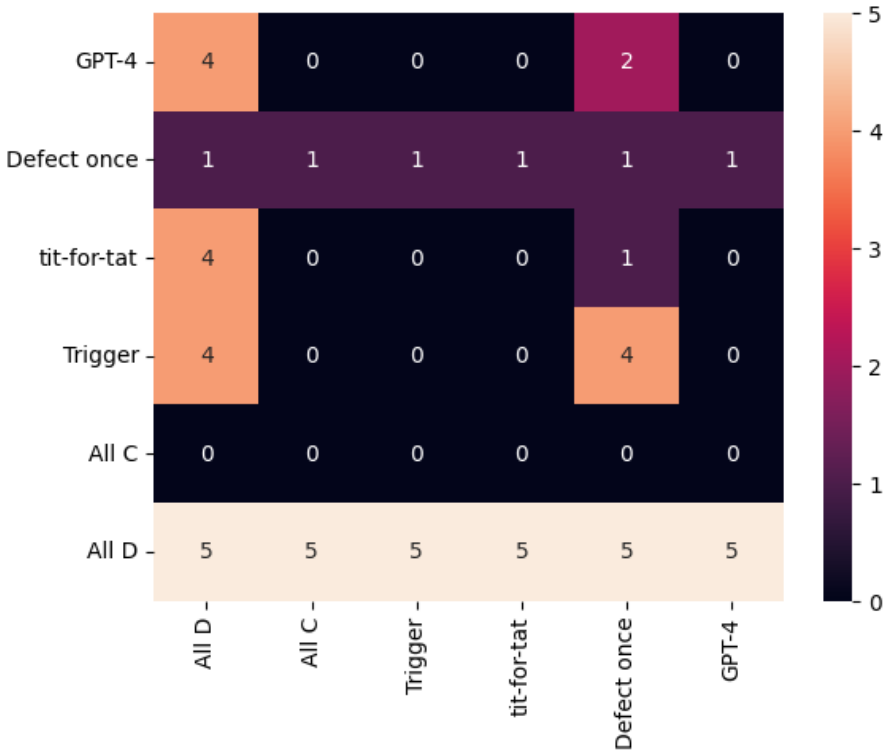

図 3: 総当たりの際の D の選択回数 (vicuna-7b-16k, ノイズ発生確率 0 %)

出してくる相手に対しての選択においては vicuna-7b-v1.5 より獲得利得が高くなるような選択をしていた. しかし, vicuna-7b-v1.5 と同様にゲームの振り返り部分に代表される応答は, GPT-4 のような高性能のモデルと比較して単調な応答となっており, 本当にゲームルールを理解して応答しているかはっきりとわかるような出力は見られなかった.

vicuna-7b-v1.5-16k の応答が怪しい例 1

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

4 今後の方針

- Llama の 7b, 13b, 70b のモデルの動作の確認
- ファインチューニング手法の調査

参考文献