

進捗報告

1 相談したいこと

- 遺伝子の表現
- 適応度の定義

2 遺伝子の表現

| 囚人A\B | C | D |
|-------|----------|----------|
| C | (p1, p1) | (p3, p2) |
| D | (p2, p3) | (p4, p4) |

図 1: 囚人のジレンマにおける利得行列

ゲームルールで変更するパラメータとして, 図 ?? の p_1, \dots, p_4 とノイズ発生確率の 5 つのパラメータがありこれらを 5 つの要素を持つ 1 次元配列として個体とする. 交差や突然変異をしつつ, $p_1 \wr p_2 \wr p_4 \wr p_3$ の大小関係を維持するような遺伝子の操作をどのように実現すればよいか.

3 適応度の定義

ゲームルールを変更することで, エージェントの振る舞いにどう影響が起こるかを見るためにエージェントのプロンプトを fix し, ゲームルールのみ変更しようと考えていた. 以前は GPT-4 エージェントがほとんど All-C となる傾向があったため, ゲームルールを変更してエージェントが D を出す傾向が多くなるようなゲームルールを得ることを目的として GA に取り組もうとしていた. 今回は, 振る舞いからペルソナを推定する方向で研究をしており, ゲームルールを変更してペルソナによって振る舞いの違いが出やすいゲームルールを得られれば良いと考えている. そういった際にはどのような指標を適用度として採用すればいいか.

参考文献