G-29

「最適」から「意図」へ：大規模言語モデルによる対戦格闘ゲーム AI の戦略理解

LIU Yiyang[*]　CHUANG Boyu[*]　THAWONMAS Ruck[**]

概要

本研究では，対戦格闘ゲームの AI 開発において，従来の強化学習が持つ「機械的」な振る舞いの限界を乗り越えるため，大規模言語モデル（LLM）の応用を試みる．2D 格闘ゲームプラットフォーム「DareFightingICE」[1] を舞台に，トップ AI「Blackmamba」の対戦ログから教師あり学習データセットを構築し，Meta の Llama 3.2 3B モデルをファインチューニングした．単語全体の埋め込みによるアクション表現の最適化，戦略的意図を説明するテキストの自動生成，そして vLLM による推論遅延の最小化といった手法を組み合わせることで，リアルタイム性を維持しつつ，戦術的文脈を深く理解する能力を持つ格闘ゲーム AI の実現可能性を示した．

キーワード

大規模言語モデル（LLM），格闘ゲーム AI，DareFightingICE，教師ありファインチューニング（SFT），Whole Word Embedding

1. はじめに

近年，格闘ゲームにおける AI の応用は多くの成果を上げており，特に強化学習（RL）を用いた手法は，多くのルールベース AI や一部の人間プレイヤーを打ち負かすエージェントを訓練可能にしている．しかし，これらの手法はアクション選択の「最適性」にのみ焦点を当てがちで，「戦術的構造」や「ゲームの文脈」をモデル化する能力に欠ける．そのアクション選択は効果的であるものの，人間プレイヤーのような戦略性や一貫性に乏しいことが多い．

また，大規模言語モデル（LLM）は言語理解，推論，マルチモーダルなシナリオにおいて強力な文脈モデリング能力を示している．そこで我々は，LLM に基づく格闘ゲーム AI を構築し，既存のトップレベル AI の行動データを学習することで，ゲームの状態と戦略を「言語レベルで理解」し，解釈可能な意思決定を出力できるかを探求することを提案する．

2. 関連研究

最近の研究，例えば LLM Colosseum [2] は，「格闘ゲーム」をプラットフォームとして LLM エージェントを評価する新たな手法を提案した．このプロジェクトでは，『ストリートファイター III』を用いて複数の主要な商用およびオープンソース LLM（例：GPT-4o，Claude，Mistral）をテストし，ELO レーティングメカニズムを通じてそのゲーム内での総合的なパフォーマンスを測定した．このアプローチは，シナリオの現実性，意思決定の複雑性，評価基準の明確さといった利点を持つ．

しかし，このプロジェクトは主に LLM 間の対戦評価に焦点を当てており，強化学習 AI や人間プレイヤーとのインタラクティブな対戦には触れておらず，モデルの特定のチューニングも行っていない．我々は，これを「LLM を訓練するための格闘エージェント開発プラットフォーム」としてではなく，ベンチマークテストプラットフォームとしての方が適していると考察する．

3. 提案手法

本研究で用いた手法の概要を以下に示す．

(1) モデル選択とデプロイ

格闘ゲームは高いリアルタイム性が要求されるため，0.6B から 4B のパラメータ数を持つオープンソースモデルの中から選定を行った．最終的に，Meta が公開した Llama 3.2 3B モデルをベースモデルとして採用し，vLLM フレームワーク [3] を用いて推論遅延を低減させた．

(2) ファインチューニング戦略

QLoRA（Quantized Low-Rank Adaptation）[4] によるファインチューニングを行う.QLoRA は, モデルの軽量化と効率的なパラメータ調整を可能にする手法であり, 比較的小規模なリソースで高いパフォーマンスを実現するのに適している. さらに、教師ありファインチューニング（SFT）手法を用いてモデルをファインチューニングした．ファインチューニングデータは，強化学習 AI「Blackmamba」が 2021 年の FightingICE 大会 [5] で複数の対戦相手（例：MctsAi23i，Thunder2021）と戦った際の対戦ログから作成した．各サンプルは以下の要素で構成される．

[*] 立命館大学情報理工学研究科
[**] 立命館大学情報理工学部

- `instruction`：固定形式の指示文
- `input`：現在のゲーム状態を記述した JSON
- `output`：実行すべきアクションとその理由

(3) アクション表現の最適化

初期実験において，モデルが `STAND_A` のようなゲームアクションを `STAND`，`_`，`A` といった複数のトークンに分解する傾向が判明した．これは理解の曖昧さや出力の不安定性を引き起こすため，Whole Word Embedding [6] を導入した．Tokenizer を修正し，各ゲームアクションが単一の単語埋め込みとして処理されるようにした．これにより，意味的な完全性が向上し，トークン数が大幅に削減され，生成遅延も低減した．

(4) 戦略的理由の生成

モデルが「なぜ特定のアクションを選択するのか」という理解を深めるため，ChatGPT 4o mini を用いて Blackmamba の行動に対する戦略的理由のテキストを生成し，解釈性のあるデータセットを作成した．合計で約 8 000 件の高品質な訓練サンプルが生成され，各出力は「アクション名\n そのアクションを使用する理由」という形式を含む．

(5) 出力制御メカニズム

必要に応じて理由の生成を切り替えるため，`max_token` パラメータを調整して出力を制御した．

- アクションのみ出力：`max_token = 1`
- アクション＋理由を出力：`max_token = 50`

4. 実験および評価

実験環境：
OS：Ubuntu 24.04.2 LTS
CPU：Intel Xeon Gold 6230
GPU：Nvidia GV100 32GB
Memory：192GB

実験 1：プレイアビリティ検証

MctsAi23i と複数回対戦させ，`max_token = 1`（アクションのみ出力）に設定し，勝率を統計した．

実験 2：解釈能力検証

`max_token = 50` に設定し，モデルがアクションの理由を正確に生成できるかを記録・分析した．

実験 3：理解能力検証

X. Li et al. [7] のプロンプトを参考にして，対戦難易度を「簡単／普通／困難」に制御し，MctsAi23i との勝率の変動をテストした．具体的なプロンプトは以下のようになる．
簡単:Please make the player feel that the game is easy.
普通: 追加プロンプトなし
困難:Please make the player feel that the game is hard.

本研究は現在進行中であるため，具体的な実験結果については発表にて報告する予定である．

5. 結論と今後の課題

本研究は，LLM に基づく格闘ゲーム AI の実現可能性を示し，特に戦略の解釈とアクション生成において良好な言語理解能力を発揮することを示した．入力形式，Tokenizer，出力ロジックのカスタマイズと最適化を通じて，Llama 3.2 3B モデルは強化学習を必要とせず，強化学習 AI の行動を模倣し，その戦略的思考を部分的に再現することができた．

今後の課題として，以下を計画している．

- 人間プレイヤーの対戦ログを導入し，人間の戦略への適応性を向上させる．
- 強化学習 AI との対戦比較を通じて，モデルの対戦強度をさらに評価する．
- マルチモーダル LLM（例：CLIP ＋ LLM）の融合を模索し，視覚とテキストの統合処理能力を向上させる．

参考文献

[1] Khan, I., Van Nguyen, T., Dai, X., Thawonmas, R., “DareFightingICE competition: A fighting game sound design and AI competition,” *2022 IEEE Conference on Games (CoG)*, pp. 478–485, 2022.

[2] OpenGenerativeAI, “LLM Colosseum: A Benchmarking Arena for Large Language Models in Fighting Games,” GitHub repository, 2024. Available: `https://github.com/OpenGenerativeAI/llm-colosseum`

[3] vLLM Team, “vLLM: A high-throughput and memory-efficient inference engine for LLMs,” GitHub repository, 2024. Available: `https://github.com/vllm-project/vllm`

[4] Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. Qlora: Efficient finetuning of quantized llms. arXiv preprint arXiv:2305.14314, 2023.

[5] FightingICE 2021 Competition Results. Available at: `https://www.ice.ci.ritsumei.ac.jp/~ftgaic/index-R21.html`.

[6] S. Geng, S. Liu, Z. Fu, Y. Ge, and Y. Zhang, “Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5),” in *Proc. of the 16th ACM Conference on Recommender Systems (RecSys ’22)*, 2022, pp. 486–497.

[7] X. Li, Y. Xia, and R. Thawonmas, “Personalized Game Difficulty with Language Models: a Preliminary Study,” in *Proc. of 2025 IEEE International Conference on Consumer Electronics (ICCE)*, 2025.