LLMagikarp: a Large Language Model agent for Pokémon Battles

Garrett Herb
herb.45@osu.edu

1 Introduction

The development of autonomous agents built off of LLM text-based perception is a new, yet highly explored field (Wang, L. et al., 2023). The broad knowledge of LLMs make them perfect for exploring new environments that can be transformed into text. By placing these LLMs into games, we are able to observe their reasoning and behavior in novel environments (Hu, S. et al., 2024). Choosing different types of games can reveal different underlying behaviors and biases of models. A tactical battle game like Pokémon can be used to evaluate a model's ability to reason in a discrete, yet enormous action space. Pokémon, with incredibly complex action space, is an ideal candidate for a game-based environment for LLM agents as it can any given state can be translated into text losslessly (Hu, S. et al., 2024).

An agent of this nature has been created with promising results with a win rate versus humans of 49% (Hu, S. et al., 2024). It translates the current and the two previous states into text and then prompts GPT-4 to generate an action. It leverages self-consistency (Wang, X. et al., 2022), or taking the most voted on action from multiple generations, to consistently make better decisions. Additionally, the PokéLLMon approach utilizes knowledge augmented generation where the state description is given facts about some of the attributes of the environment so the agent can make more informed decisions.

With the tools at the battler's disposal, its results are promising, but there is much room for improvement. The following augmentations are meant to not only improve the models baseline performance, but also allow the model to refine its battling over time without parameter updating: (1) A smaller, open-source model will be used to compare performance with GPT-4; the model will also be fine-tuned to instill Pokémon specific knowledge into it which could (i) improve action generation (Liu, S. et al., 2024) and (ii) reduce overhead by making the knowledge augmented generation obsolete. (2) Add a skill library that can enable lifelong learning through reflection of successful and unsuccessful battles (Wang, G. et al., 2023). (3) Employ the ability to think of what the opposition might do and act according to that information. (4 if time permits) Equip the agent with battling specific tools that can drastically increase the specificity in which it makes decisions.

2 Pokémon

The popular franchise Pokémon centers around a series of video games where players explore a fantasy world where monsters, or Pokémon, are able to be captured and trained. To progress in the mainline games, the player must battle other 'trainers.' This battling mechanism has become a huge cultural sensation with many events pinning players against each other for formal competitions. While the premise of the battle is simple, eliminate your opponent's Pokémon before your own are eliminated, the battling is extremely complex and requires players to consider a huge amount of information and possibilities.

While many battling formats exist, the agent will only be employed in random single battles. This adheres to the traditional single battle format of the Pokémon games while exposing the agent to much more diverse starting states from the random aspect. In random single battles, each player is given 6 random Pokémon, where only one can be active at each time. For each turn, both players are able to choose from a couple action choices. They can either choose one of 4 moves the active Pokémon has access to or switch the active Pokémon with an unactive Pokémon that has not been knocked out yet. Both players pick actions in order to reduce the health of the opposing Pokémon to knock them out and to keep their own Pokémon from accruing damage. Again, this premise seems simple, but there are many variables that contribute to a game where the next state is impossible to consistently predict.

2.1 Pokémon species

Each of the 6 random Pokémon on a given team are one of 1,025 unique species all with attributes that pertain to that specific species (Bulbapedia, 2024b).

2.2 Moves

In a battle, all Pokémon have 4 moves. Moves are meant to either damage a targeted Pokémon, inflict a negative status to a target Pokémon, inflict a positive status on the targeted Pokémon, or heal a targeted Pokémon. While there are only 4 moves per Pokémon in the battles, there exist 934 unique moves that Pokémon are able to learn (Bulbapedia, 2024b).

2.3 Types

There exist 18 elemental types where each Pokémon can have either 1 type or a combination of 2 types. Each move has a single elemental type as well. Every type has certain weaknesses and advantages against other types (Bulbapedia, 2024b).

2.4 Abilities

Every Pokémon has 1 to 4 innate 'abilities' that can alter the effect of opposing moves on itself or affect its own attacking moves (Bulbapedia, 2024b).

2.5 Stats and Nature

All Pokémon have 6 stat categories (Health, Attack, Defense, Special Attack, Special Defense, and Speed). Each stat can be balanced to allow Pokémon to be faster, stronger attackers, stronger defenders, etc. All Pokémon have their own initial spread of these stats that can also be augmented by Pokémon natures. Each Pokémon can be one of 25 natures that correspond to a change in initial stats (Bulbapedia, 2024b).

3 Method

3.1 LLMs agents for Games

The PokéLLMon show promise when utilizing a LLM for action generation in the Pokémon battling system (Hu, S. et al., 2024); they also demonstrate the success of equipping a model with game specific knowledge directly into state prompts. (Costarelli, A. et al., 2024) demonstrate that even when equipped with chain of thought (CoT) prompting (Wei, J. et al., 2022) or reasoning via planning (RAP) (Hao, S. et al., 2023), LLMs in strategic game-based environments still are outperformed by humans are lacking some type of the reasoning process behind effective strategy creation where a prompt to think of an opposition's strategy could help. The use of a smaller model can prove to be hurtful when high level reasoning and planning is needed, however, Liu, S. et al. (2024) has demonstrated that for generating actions in game environments, small LLMs that are fine-tuned on game specific knowledge can outperform bigger flagship models like GPT-4.

3.2 Lifelong Learning for LLM agents

(Shi, H. et al., 2024) discuss the need for continuous learning to combat catastrophic forgetting and how retrieval augmented generation (RAG)-like (Lewis, P. et al., 2020) can be applied to aid this. Lifelong learning for agents in game-based environments has been explored by Wang, G. et al., (2023) where they displayed that a skill library that can be called upon when generating further action enables agents to explore more of a vast action space and also have temporal progression without the need to update model parameters.

4 Method

4.1 Open-source model adaptation and fine-tuning implementation

Llama 3.2 3B and Llama 3.1 8B will be implemented into the PokéLLMon framework to explore the effect of SoTA small models affect the agent's performance against humans. Additionally, the better of the two models on the initial run of the program will undergo a low-rank adaptation (LoRA) (Hu, J.E. et al., 2021) fine-tuning on Pokémon battling specific facts.

4.2 Skill library implementation

After each game, the agent will be asked to review its game and give itself high-level feedback upon what went well or went poorly. The feedback will then be stored where it can later be accessed when faced with a similar battling situation.

4.3 Oppositional thinking implementation

Before each state is created, the agent will attempt to reason what the opposition might do on the given turn. This response will then be appended to the current state for action generation.

4.4 Battle specific tools implementation

Human players commonly use damage calculator tools to determine the damage a move might do on an opposing Pokémon. This insight can give the model more confidence and precision when choosing its next moves or switches. Additionally, giving the agent access to a database to lookup potential moves/abilities/stats for a given Pokémon could help its predictive power.

4.5 Result analysis

The effect of the previously mentioned augmentations on model effectiveness will be determined by its performance against a heuristic bot created by Hu, S. et al., (2024) and against human players. The win rate as a percentage will be used as the evaluation metric. Additionally, the win rate overtime will be analyzed to determine the effectiveness of the skill library.

5 Anticipated results

It is expected that the smaller models will initially not be as capable as the larger GPT-4, but once fine-tuned, they should be comparable if not outright better.

The implementation of the skill library is expected to increase the performance of the agent as plays more and more battles against a diverse host of scenarios.

The oppositional thinking should help with reasoning and ultimately improve performance against capable players assuming the agent can make reasonable predictions; however, against unexperienced players, it could result in misplays as the agent would assume predictable behavior where an unexperienced player could play unpredictably.

Giving the model battle specific tools could drastically increase the performance and specificity of its moves as it can mathematically justify actions versus arbitrarily making them on hunches.

References

Bulbapedia. 2024b. List of abilites. September. https://bulbapedia.bulbagarden.net/wiki/Ability#List_of_Abilities

Bulbapedia. 2024b. List of moves. September. https://bulbapedia.bulbagarden.net/wiki/List_of_moves

Bulbapedia, 2024b. List of Pokémon by National Pokédex number. September. https://bulbapedia.bulbagarden.net/wiki/List_of_Pok%C3%A9mon_by_National_Pok%C3%A9dex_number

Bulbapedia, 2024b. List of stats. September. https://bulbapedia.bulbagarden.net/wiki/Stat

Costarelli, A., Allen, M., Hauksson, R., Sodunke, G., Hariharan, S., Cheng, C., Li, W., & Yadav, A. (2024). GameBench: Evaluating Strategic Reasoning Abilities of LLM Agents. *ArXiv, abs/2406.06613*.

Hao, S., Gu, Y., Ma, H., Hong, J.J., Wang, Z., Wang, D.Z., & Hu, Z. (2023). Reasoning with Language Model is Planning with World Model. *ArXiv, abs/2305.14992*.

Hu, J.E., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., & Chen, W. (2021). LoRA: Low-Rank Adaptation of Large Language Models. *ArXiv, abs/2106.09685*.

Hu, S., Huang, T., Ilhan, F., Tekin, S.F., Liu, G., Kompella, R.R., & Liu, L. (2024). A Survey on Large Language Model-Based Game Agents. *ArXiv, abs/2404.02039*.

Hu, S., Huang, T., & Liu, L. (2024). PokeLLMon: A Human-Parity Agent for Pokemon Battles with Large Language Models. *ArXiv, abs/2402.01118*.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Kuttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *ArXiv, abs/2005.11401*.

Liu, S., Li, Y., Zhang, K., Cui, Z., Fang, W., Zheng, Y., Zheng, T., & Song, M. (2024). Odyssey: Empowering Agents with Open-World Skills. *ArXiv, abs/2407.15325*.

Shi, H., Xu, Z., Wang, H., Qin, W., Wang, W., Wang, Y., & Wang, H. (2024). Continual Learning of Large Language Models: A Comprehensive Survey. *ArXiv, abs/2404.16789*.

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W.X., Wei, Z., & Wen, J. (2023). A Survey on Large Language Model based Autonomous Agents. *ArXiv, abs/2308.11432*.

Wang, X., Wei, J., Schuurmans, D., Le, Q., Chi, E.H., & Zhou, D. (2022). Self-Consistency Improves Chain of Thought Reasoning in Language Models. *ArXiv, abs/2203.11171*.

Wang, G., Xie, Y., Jiang, Y., Mandlekar, A., Xiao, C., Zhu, Y., Fan, L., & Anandkumar, A. (2023). Voyager: An Open-Ended Embodied Agent with Large Language Models. *Trans. Mach. Learn. Res., 2024*.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Chi, E.H., Xia, F., Le, Q., & Zhou, D. (2022). Chain of Thought Prompting Elicits Reasoning in Large Language Models. *ArXiv, abs/2201.11903*.

Zhang, Y., Mao, S., Ge, T., Wang, X., Wynter, A.D., Xia, Y., Wu, W., Song, T., Lan, M., & Wei, F. (2024). LLM as a Mastermind: A Survey of Strategic Reasoning with Large Language Models. *ArXiv, abs/2404.01230*.