Adapting Visual LLMs for gameplay in Pokémon FireRed

Some thoughts on previous works.

By Issac Lee, Erica Wu, Leo Mai

Why do we care about solving Pokémon Red?

Pokémon Red is a Hard Problem

- How do we keep track of what we seen?
- How do we know what a “good” move is?
- How can the LLM use natural language in a non explicitly natural language problem?

Dataset

- *Domain Specific?*
- *Clean?*

LLM

- *Finetuned?*
- *RLHF?*
- *Architecture?*

Interaction

- *Tool Calling?*
- *RL?*


LLM Based
LLM + Fine-Tuning
LLM + Tool Calling
LLM Guided RL
RL Based
Just throw a LLM at the problem
Train it with some domain knowledge.
Give it a MCP server. Moves some logic into predefined tools.
Use LLM to control RL instead of a predefined reward function..
Just throw RL at the problem.

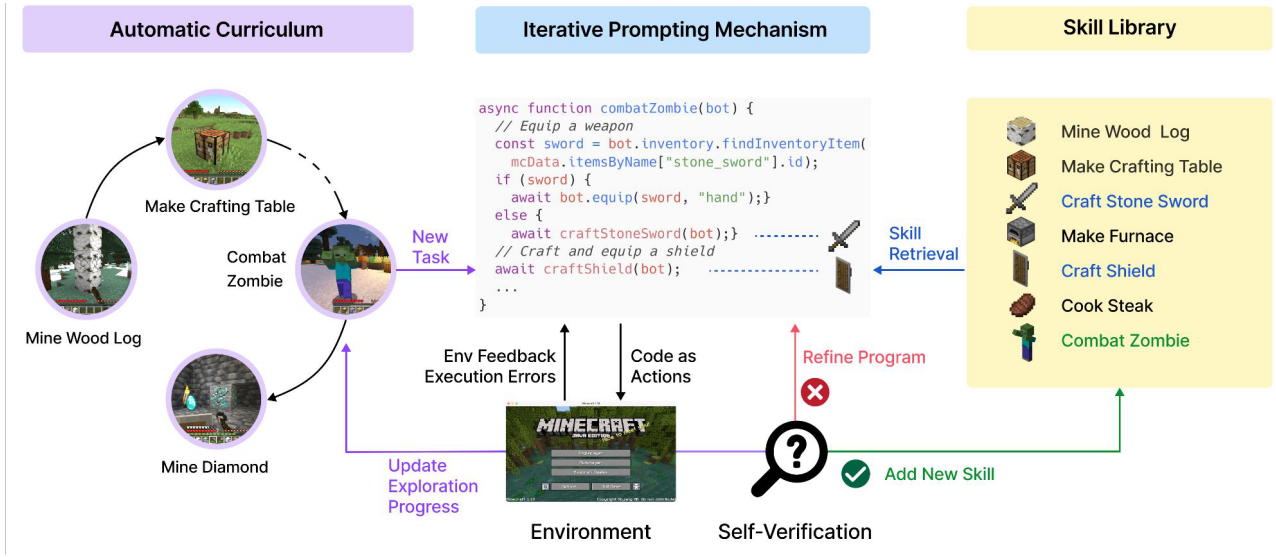
Previous Examples

LLM Based

- Pure Prompting
- ChatGPT 3 (2022)
- Context Window Only

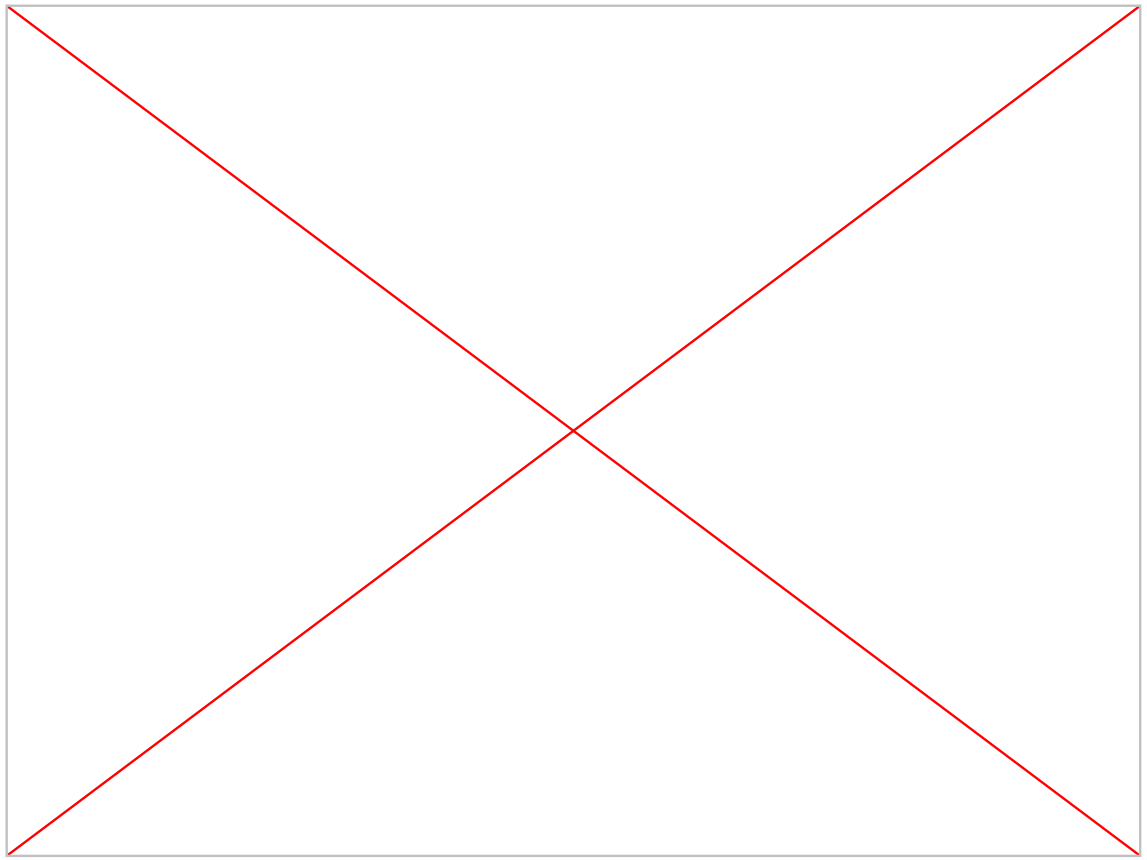
LLM + Finetuning

- Andy 3.6
- Q&A Dataset
- Uses special commands as output



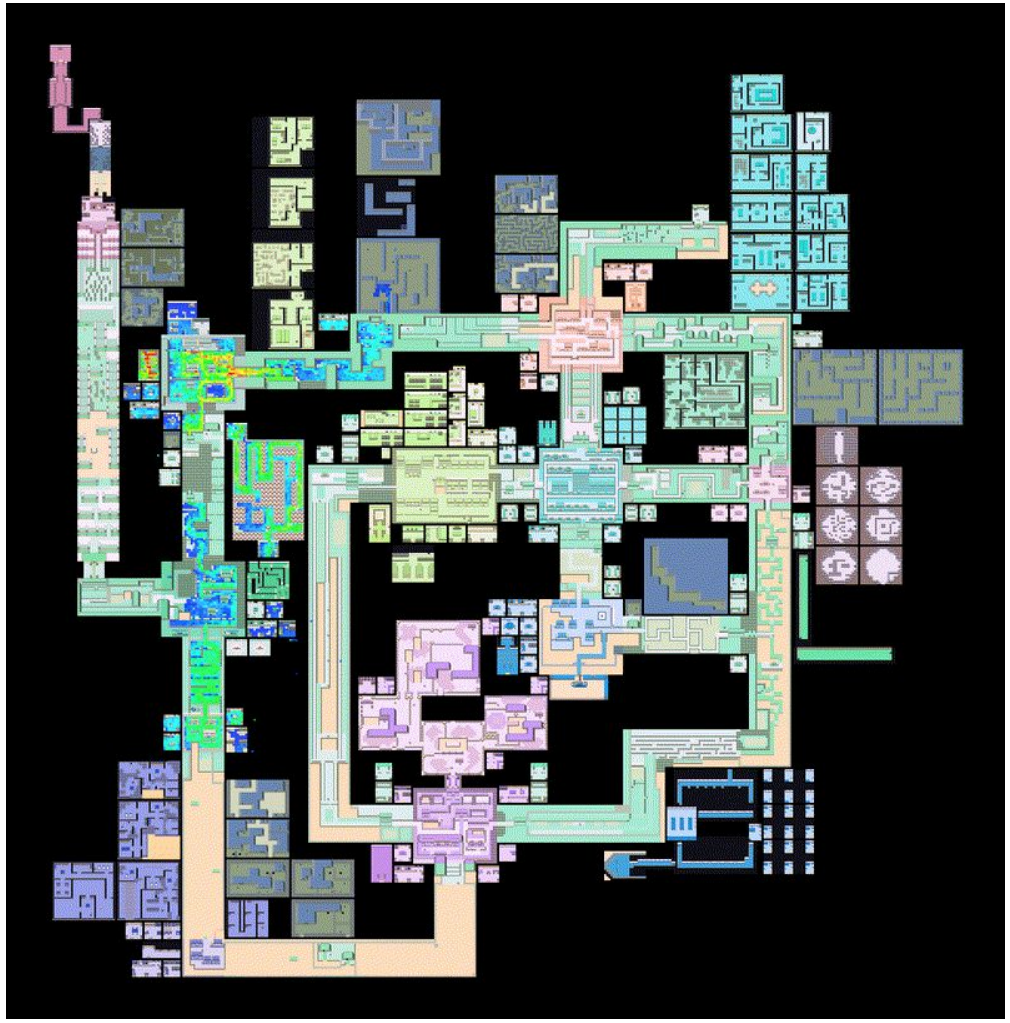
LLM + Toolcalling

- Skill library of previously learned tasks
- RL like learning environment
- Writes its own code per skill



LLM Guided RL

- Hard to create reward functions
- LLMs good at steps by step RL good at fine details
- LLMs help transform problems into a natural language one



Reinforcement Learning

- RL good at solving complicated tasks
- Hard to define what a good move is
- Used a lot in industry!

Main Takeaways

- Most LLMs use a mixture of these solutions to solve problems
- LLMs are good at planning, bad at logic
- Solving Pokémon Red requires a lot of the same solutions seen throughout

Questions?