Project Title

Reinforcement Learning Agent for Street Fighter II using PPO in Gym-Retro

Submitted By:

Ijlal Iqbal (22K-5034)

Huzefa Saifuddin (22K-5125)

Areeb ur Rehman (22K-6003)

Course:

AI

Instructor:

Alishba Subhani

Submission Date:

May 11th, 2025

1. **Executive Summary**

Project Overview:

This project aimed to train an AI agent to play Street Fighter II using reinforcement learning, specifically the PPO algorithm. By utilizing the Gym-Retro environment, the project captured game frames, defined a custom reward function, and allowed the agent to learn by playing against a fixed AI opponent. Modifications include simplifying the action space and introducing reward shaping to encourage strategic behavior and improve learning efficiency.

2. Introduction

Background:

Street Fighter II is a popular 2D arcade fighting game involving two characters engaging in timed combat using special moves, blocks, and attacks. This project chose Street Fighter II because it offers a fast-paced environment with real-time decision making. To adapt it for AI learning, we used Gym-Retro to emulate the game, simplified the control space, and created a reward function to reflect meaningful combat behavior.

Objectives of the Project:

- Train an RL agent to effectively play Street Fighter II against a fixed AI.
- Design a reward function to encourage winning strategies.
- Use PPO to optimize agent policies over time.
- Evaluate performance in terms of win rate and gameplay quality.

3. Game Description

Original Game Rules:

Street Fighter II is played between two players who select characters with unique move sets. The objective is to reduce the opponent's health bar to zero within the round timer. Each match consists of up to three rounds. Players can move, jump, block, and perform light/heavy attacks.

Innovations and Modifications:

- Limited the character to Ryu for training simplicity.
- Fixed the opponent character (Ken) for consistent evaluation.

- Implemented frame skipping to speed up training.
- Reward shaping introduced to promote strategic actions (e.g., damaging opponent, avoiding hits).

4. AI Approach and Methodology

AI Techniques Used:

- Reinforcement Learning using PPO.
- Neural network policy models trained using Stable-Baselines3.
- Gym-Retro as the emulation environment.

Algorithms and Heuristic Design:

- Positive reward for damaging the opponent.
- Negative reward for taking damage.
- Large bonus for winning a round.
- Penalty for inactivity to prevent stalling.

Heuristics were designed to balance aggression and defense.

AI Performance Evaluation:

Performance was measured using win rate over episodes, health remaining, and reward per episode. Training graphs showed steady improvement with PPO. Final models consistently beat the fixed AI.

5. Game Mechanics and Rules

Modified Game Rules:

- One playable character: Ryu.
- Fixed opponent: Ken.
- Training limited to one stage.
- Best of 3 round format.
- Frame skip set to 4 to optimize processing.

Turn-based Mechanics:

The game runs in continuous time, with the AI selecting actions at every frame (after frame skip).

Winning Conditions:

The agent wins a match by defeating the opponent in two out of three rounds by reducing their health to zero.

6. Implementation and Development

Development Process:

The project began with setting up the Gym-Retro environment and extracting meaningful observations from screen frames. A custom wrapper was written to define the action and observation space. PPO from Stable-Baselines3 was used to train the agent. Reward functions were iteratively refined. Training progress was monitored via tensorboard and periodic evaluations.

Programming Languages and Tools:

- Programming Language: Python
- Libraries: Gym-Retro, Stable-Baselines3, NumPy, OpenCV, PyTorch
- Tools: GitHub for version control, TensorBoard for training visualization

Challenges Encountered:

- Complex frame data requiring preprocessing.
- Sparse rewards leading to slow early learning.
- Balancing exploration vs exploitation in PPO.
- Managing long training times and stability.

7. Team Contributions

Ijlal Iqbal (22K-5034): Responsible for AI algorithm development and PPO training.
Huzefa Saifuddin (22K-5125): Handled Gym-Retro setup, action space design, and training supervision.
Areeb ur Rehman (22K-6003): Focused on environment wrapping, reward design, and evaluation reporting.

8. Results and Discussion

The trained PPO agent achieved a win rate of approximately 80% against the fixed AI opponent. Decision times were acceptable for real-time performance, and the learned policy showed intelligent behavior such as avoiding attacks, corner trapping the opponent, and timing special moves. Training progress showed consistent improvement in total reward per episode.

9. References

- Sutton & Barto - Reinforcement Learning: An Introduction
- Gym-Retro Documentation
- Stable-Baselines3 Documentation
- PPO Algorithm Papers (OpenAI)
- Online tutorials and GitHub repositories related to RL in games