IL RE BOMBA

Improving Pokémon AI Traning With NSGA-II

Simone Alghisi Samuele Bortolotti Massimo Rizzoli
Erich Robbi

University of Trento

August 23, 2022

Introduction

1. introducing pokemon, and RL in a few words
2. describe why the training becomes very difficult (THE NUMBERS MASON, WHAT DO THEY MEAN)
3. propose the solution with NSGA-II in order to have a controlled search rather than a completely random one
4. describe the multi-objective problem:
    - genotype representation
    - mutation and recombination strategy
    - search strategy
    - defining objectives and the optimisation
5. specify what kind of tests have been conducted, and why (IDK EITHER il re bomba)
6. live demo
7. analysis of the results
    - pareto front (we show both plots of the same pareto front)
    - training convergence
8. difficulties
9. contributions

UNIVERSITÀ DI TRENTO

Introduction

Pokémon

Pokémon uses a turn-based system: at the start of each turn, both sides can choose to attack, use an item, switch the Pokémon for another in their party. The Pokemon who strikes first is determined by the Move's Priority and the Pokémon Speed. Each Pokémon uses moves to reduce their opponent's HP until one of them faints, i.e. their HP reach 0. If all of a player's Pokémon faint, the player loses the battle.

Figure 1: Pokémon battle

Reinforcement Learning

Reinforcement learning (RL) is an area of Machine Learning where an agent receives a reward based on the action it has performed. Actions allow the agent to transition from a state to another. The final objective is to learn a policy to reach a terminal state with the best reward achievable.

Deep Q-Learning

The reinforcement learning technique we have employed is called *Deep Q-Learning*, which maps input states to a pair of actions and Q-values using an Artificial Neural Network. *Q-Learning* is based on the *Q-function*, namely $Q : S \times A \rightarrow R$, which returns - given a state-action pair $(s, a \in S \times A)$ - the expected discounted reward $(r \in R)$ for future states.

NSGA-II

NSGA-II is a Evolutionary Algorithm that allows to produce *Pareto-equivalent* (or non-dominated) solutions of a multi-objective optimisation problem.

General idea

The idea is that, given that the search space is very big - there are 10^{354} different ways a Pokémon battle can start, and each turn has at most 306 different outcomes (and only for a single player) - we would like to positively bias our model with a controlled search, removing particularly useless moves, i.e. consider for the most Pareto-equivalent solutions.

Genotype representation

Generally, in a Pokémon battle two actions are possible, i.e. performing a move or a switch. Moreover, depending on the type of battle, it may be necessary to specify the target of the move. To encode such a thing, we came up with the following genotype: each Pokémon is represented using two genes, i.e. action and target (optional) (a, t). The whole genotype tells us who is going to perform what on whom.

Genetic operators

Mutation

Mutation is performed for each gene in a genotype with probability $\mathbb{P}_m = 10\%$: both the action and the target may be mutated, meaning that it is possible to go from a move to a switch (and vice-versa).

Recombination

Instead, we used Uniform Crossover in a particular way: given that each Pokémon is represented by a valid (a, t) pair, we perform crossover by selecting the whole pair from one of the parents to avoid inconsistencies. Furthermore, crossover is performed with $\mathbb{P}_c = 100\%$, and $\mathbb{P}_{bias} = 50\%$ (i.e. the bias towards a certain offspring).

Search strategy

Objective & Optimisation

Il Sium del Re Bomba

A

B

Live Demo

Gif in PDF

Resources

Repositories

- pareto-epsilon-greedy-RL
- poke-env (modified)
- Pokemon_info

Collaborators' Github

- Simone Alghisi
- Samuele Bortolotti
- Massimo Rizzoli
- Erich Robbi

Thanks for your attention!

References I