Proposal - Playing pokemon battles optimally

Julian Schubert

November 12, 2021

1 Introduction

Pokemon is a video game series created by the Pokemon Company. The goal of the game is to not only catch pokemon, but also to train them and use them to battle other trainers. In the mainline games, the focus lies on the story as well as exploring the world. The pokemon genere has evolved quite a bit since it's early days, Figures 1 and 2 contain screenshots of the first pokemon games, pokemon red and pokemon green. Figures 3 and 4 contain ingame footage of

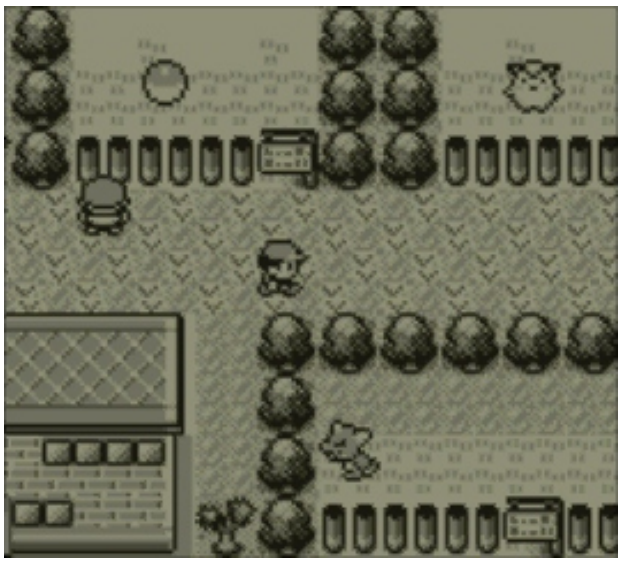

Figure 1: Exoploring the map in pokemon red

Figure 2: Fighting another trainer in pokemon red

Image source: nintendo.de

the latest games in the series, pokemon sword and shield. This thesis will exclusivly focus exclusivly on battling as there detailed lists of the locations of everything there is to explore within the games.

Figure 3: Exoploring the map in pokemon sword.
Image source: pokemon.com

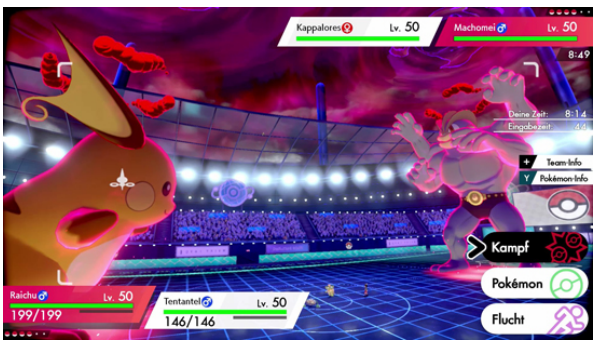

Figure 4: Fighting another trainer in pokemon sword
Image source: Image source: nintendo.de

2 Prerequisites

Nintendo does not provide an API for the game, however, the fan project Pokemon Showdown is a free online tool that can be used to battle online trainers. On top of that, it provides the functionality to use their simulator offline in an CLI. The entire source code for pokemon showdown can be found at Github. Additionally, the python library poke-env is used as it provides a convenient interface to showdown.
The combination of both tools allows to easily perform deep reinforcement learning on a local running instance of showdown, as well as evaluating models against human players on the official pokemon showdown server. Figure 5 shows a battle between two human players on pokemon showdown. Lastly, the creators of pokemon showdown provided me with over 8 Million replays of games played by humans.

3 Battling

The focus of this thesis will be on batteling. Each player has a team of six pokemon. The pokemon are chosen randomly from the available pokemon in the game. Currently, there are 898 different Pokemon, some of them even have different forms. Each pokemon knows 4 different, also randomly assigned, moves. However, not every Pokemon can learn every move.
A move can either damage the opposing pokemon, heal the own pokemon, increase offence or defence, or add status effects like paralysis or sleep.
The game works based on a rock-paper-scissors-like system. Each pokemon has one or two types, each move has a type as well. Fire-Type moves are strong against Grass types, Grass types are strong agains Water pokemon and water pokemon are strong against fire pokemon. In total, there are 18 different types.
There are other aspects like weather, speed, base stats and level which heavily influence the outcome of a battle, they won't be discussed in this proposal, however they will have to be taken into account in the thesis.

3.1 Advanced battling strategies

In section 6 a simpe rulebased approach is introduced that always picks the move that deals the most amount of damage. However, picking an optimal move is a lot more complex than this. Pokemon battles are not just strict one on one battles where the player with the most damage dealing pokemon always wins. In competetive play, a wide variety of non- or low-damaing moves is used. Dragon Dance does not deal any damage to the opponent, but increases the damage the pokemon deals in the next moves.
Additionally, each pokemon can play one or more different roles where the playstyles, again, counter each other in a rock paper scissors like fashion. Identifying in a random team what playstyle to pick for which pokemon is essential to winning battles. This decision is also based on the playstyles the opponent chooses.
Lastly, predicting enemy decisions is a key component of competive play. For example, if the pokemon of player 1 is strong against the pokemon of player 2, the opposing player will likely switch to another pokemon. As an anticipation of this move, the first player could pick a move that will deal less damage against the current opposing pokemon but is also good against the pokemon he assumes the enemy to switch to.
These elements turn competetive Pokemon in a game where traditional stratagy and probability management are combined with information management.

4 Execution

This thesis will investigate multiple possible approaches to optimize battling. The first approach will be a rule based. Different complexity levels will be tested against each other.
Secondly, backpropagation will be used to train a neural network to play like a human player. The traning data was provided to me by the pokemon showdown team, it contains over 8 million replays of random pokemon battles.
Lastly, a reinforcement algorithm will be tested. The possibility to pretrain the network using either rules or replay data will be investigated as well.

5 Evaluation

The poke-env libary provides not only a random player, but also a max damage player that always chooses the move with the highest base damage. A simple reinforcement approach is given as well. These three agents will be used as baseline.

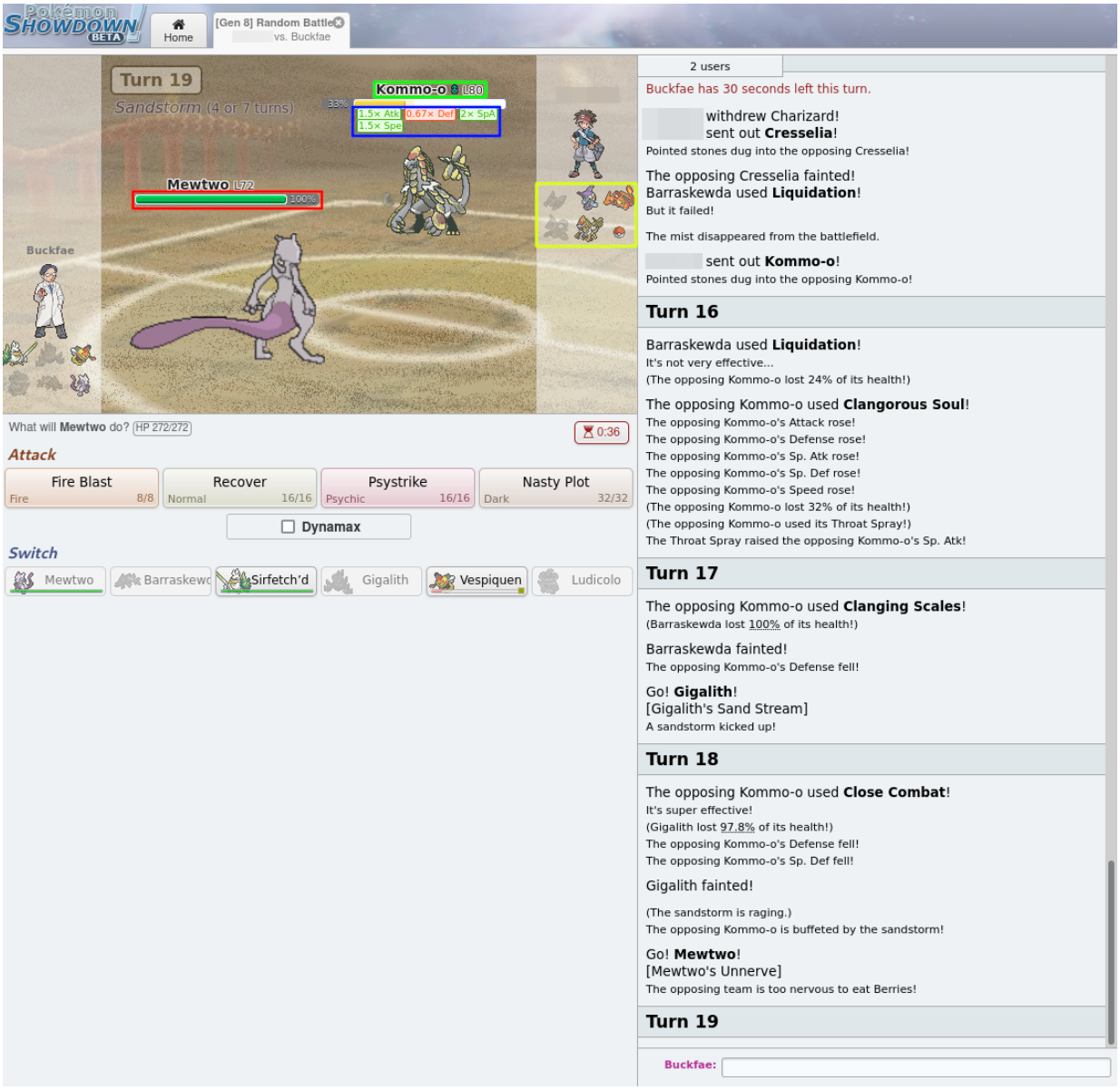

Figure 5: Screenshot of me batteling a random opponent

Pokemon Showdown also has an Elo-System similar to chess. The authors of Showdown allow bots to compete in ranked matches, so the approaches developed in this thesis will also be evaluated by playing ranked games against actual humans.

6 Example rulebased approach

During testing of the techincal feasibility of this approach, I developed a simple rule based approach.

6.1 Damage Calculation

The expected damage a move will deal to an opposing pokemon is calculated as follows:

$$\text{Expected Damage} = \text{Move Base Damage} \times \text{Move Type Modifier} \times \text{Stab Modifier} \times \text{Move Accuracy}$$

Stab Modifier: If the type of the move is equal to one of the possibly two types of the pokemon, the move will deal 1.5 times more damage.

6.2 Rules for playing

Deciding on the next move is purely based on the expected damage in the current turn. If a pokemon faints, the next pokemon to switch in is picked on based on the expected damage it can deal to the opposing pokemon.

6.3 Evaluation

This approach was evaluated against a random player and a player that always chooses the move with the hightest base damage. In 500 games, this approach won 484 out of 500 games against the random player and 419 / 500 games against the player that always chooses the move with the highest base damage.

7 Time Management

Following table shows the general timeline for this thesis.

	Time in weeks									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Mechanics of the game	x									
Extracting Data from replays		x								
Rulebased approach: Implementation			x	x						
Buffer					x					
Reinforcement learning: Research						x				
Reinforcement learning: Implementation							x	x		
Writing	x	x	x	x	x	x	x	x	x	x

Table 1: Time Management in weeks