Assignment3

Yann Amado, Micky Labrèche

2022-12-08

1 Improving a figure

The document we have chosen is:

Deep Reinforcement Learning in Atari 2600 Games by **Daniel Bick**

(https://fse.studenttheses.ub.rug.nl/20812/1/AI_BA_2019_DANIELBICK.pdf)

The thesis concerns itself with comparing two reinforcement learning algorithms on their performances of playing different atari games. The compared algorithms are Deep Q-learning (DQN), and Deep Quality-Value Learning (DQV). Different strategies within these are compared too (NoisyNet- and diversity-driven approaches for both algorithms distributively).

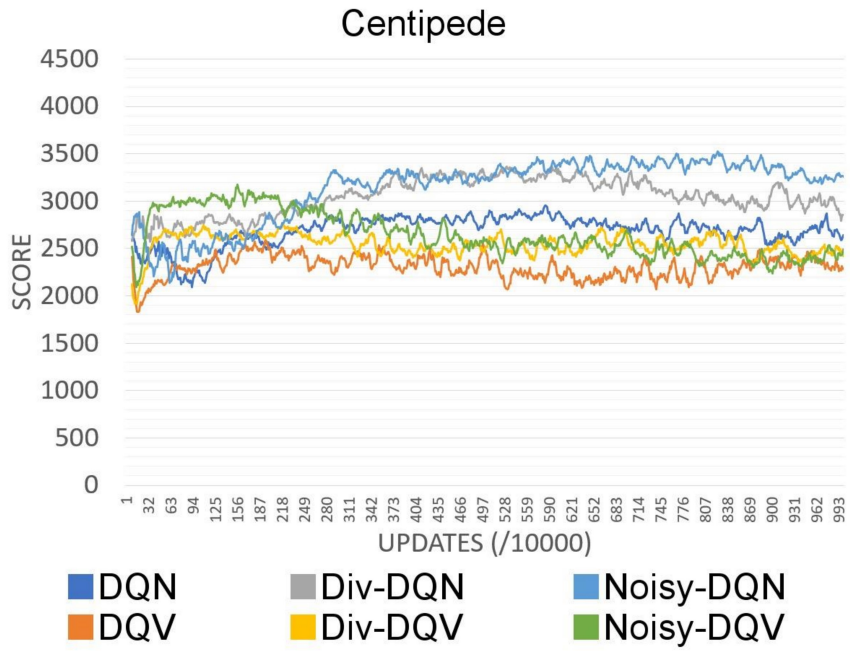
The chosen plot for this assignment (Figure 5.2 in the thesis) shows the averaged performance of each of the six versions of the algorithms for a specific game called centipede. The Y axis shows the performance of the algorithm in terms of points obtained in the game. The X axis shows the amount of updates at the moment of the score.

We will try to improve this visualization by applying the design principles for graphics:

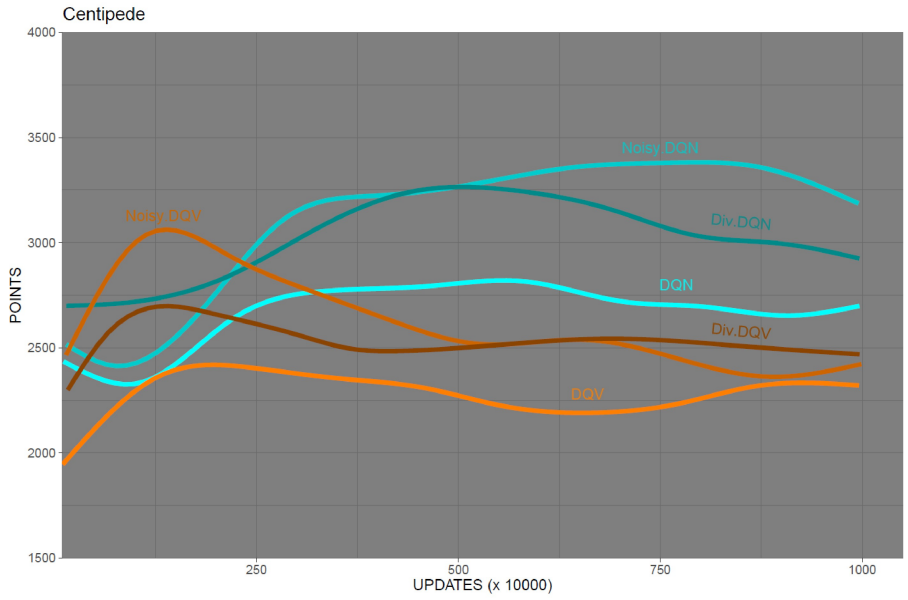
1. Focus what you want to communicate: For this graph, the importance lies on the difference in performances between the groups of algorithms. Because of this, it's important that the plot focuses on the differences between the groups themselves and each algorithm between each other; hence, we believe to be better to smooth out the lines, since the focus is *between* the algorithms and not between individual runs, making it easier to discern between them.
2. Facilitate identification: Taking into account the text above the graph, this section mentions that the three DQN implementations outperform the DQV implementations in the last two-thirds. Although this is observable, it's not very clear. With this in mind, we clustered the lines in groups by using different colors between groups, but only different color gradients of the same color between the group. With this manner, it's possible to instantly notice to which group a line belongs to.
3. Account how we perceive quantity: Even though starting the axes at 0 is generally a good practice, we opted to start the y axis at 1500 for a couple of reasons: first, having the graph start at 0 gives an effect of "zoom out", making the important part of the plot (the one that contains the lines) smaller and reducing clarity; second, since the focus is on the difference between algorithms, the user will need to look at the y-axis for both lines if he wants to quantify the distance between them, making it impossible to miss the altered scale and taking away the need for it to start at 0. This is a peculiarity of this graph and dataset specifically.

Another change we made was to take out the irregular ticks at the x axis and change into a more regular and easy to comprehend representation, reducing cluttering of the plot and making it easier to comprehend, as well as implementing vertical ticks for the reader to easily relate the x and y axis. 4. Color, order & position express information: Lastly, we chose to implement annotated graphs with the names of the algorithms instead of a legend, making it easier to read.

Old graph from paper



Improved graph



2 Pokémon graph: additional information:

In this graph we can find the worst matchup to a Pokémon. This is very practical for countering your opponent's strategies.

We can use the shiny graph to input the name of a Pokémon either via text, or from a dropdown menu. After this we can see in the graph to right the worst matchups to that input, with every dot being an individual Pokémon.

We can switch the input and see that the dots change, Pokémons have different types and therefore are more vulnerable to other certain Pokémons. We represented the types by the color of the dots, with those colors representing the actual colors used in the game. If you zoom out with the damage slider, you can see that a single type exists that is generally stronger against the selected Pokémon.

When choosing a Pokémon for a fight, the damage is very important. In this graph, we show a ranking based on "damage score", which is a score based on individual properties of a Pokémon. This is not the actual damage dealt in a fight because there can be other circumstantial factors that influence that damage.

It is also important to watch out for good defense, this is a measure of how much a Pokémon can withstand attacks, and thus, how many turns it'll partake in the fight. To measure this, we plotted the x-axis with the defense points.

Speed of attack is also an important factor to consider, and we can see this in the graph by looking at the size of the dots. When the speed is low, the Pokémon will never make the first move, which is not benefitial to him.

Lastly the name of certain Pokémon is shown if they are outlier from the current range selection. You probably want to select that Pokémon!

3 Reflection

Shiny graphs are a great tool to use if the data benefits from a more interactive approach. This interactivity is what sets aside shiny graphs from standard ones by giving a whole new set of possibilities in how we communicate the data to the user, making it possible for them to change the graph at will; this is specially important in cases where the data is very packed with information or if every single point is a potential point of interest, since the reader will now be able to analyse the portion of the data that is closer to his reality, as is the case for the Pokémon graph considering a player will most likely only have a few of the hundreds of options available. In other words, shiny graphs makes it possible to remove several different graphs and replace it for just one interactive graph, as well as giving the possibility for the reader to be able to focus on the parts of the data that it wants, instead of those pre selected by the author, making it more interesting and giving a deeper insight. If this was not possible, for the Pokémon chart to be useful, it would need to contain several different portions of the same graph, focusing on different plots, and even them it would not be possible to represent the same range of options available with the shiny graphs. However, even though the benefits are very attractive, it's use should be saved for when it's actually needed since it adds a layer of complexity for the user and can make erroneous interpretations easier to occur, like in cases where the Simpson's Paradox may be present; with this being one possible reason to restrict the user's freedom in what he can do with the graph. Nevertheless, it's a great tool for data visualization and should be properly utilized when the situation arises.

4 Contributions

Both autours worked on every item of this assignment, with the workload divided equally