**Project Proposal – Reinforcement Learning – Visualization of eligibility traces**

**Team members:** Abdelrahman Amer, Chidambar Joshi, Fida Cherni

* The team wants to present the explanation and visualization first followed by the math to understand the concept before applying it to the math.
* A screenshot of a computer

  Description automatically generatedWe will start with a transition from TD(0) and MC, mention briefly their limitations, then present the middle ground or generalization which is TD(λ).
* The team will continue to develop the heat map we created (shown on the right with part of the code) showing an agent traversing a grid world which colors the steps that the agent takes (in blue) and highlights the cells most recently visited brightly, slowly dimming the colors over time unless the state is visited again. The team wants to add a controllable time slider so the student can slide through the steps of the episode from beginning to end.
* A screenshot of a computer code

  Description automatically generatedThe colab code will test the agent’s performance in the environment by changing the following hyper parameters. (at least covering these 4 combinations, high-high, low-low, high-low, low-high)
  + λ (Trace decay factor)
  + 𝛾 (Discount factor)
* Plots will also be included to show the accumulated reward (y-axis) over the number of episodes (x-axis) to compare the performance of different agents for the different hyper-parameters selected above (including for MC and 1-step TD).
* The team also wants to find real world applications that really would benefit from this technique and present them with images in the colab (even if there is no code to run).
* A diagram of a diagram

  Description automatically generated with medium confidenceIf the colab is not too long by this point, we also want to include different ways to compute the z value (eligibility trace) as shown on the right.
* Additionally, we can add a brief discussion of the limitations of this method.
* Generally, the team would like to gamify the presentation as much as possible, for example by emphasizing the WINNING agent and focusing on their characteristics (hyperparameters), counting the NUMBER OF STEPS in each episode to see the FIRST agent that reaches the destination in the lowest number of steps attributing that win to the correct allocation of rewards to previous steps visited and how the most recent steps contributed the most to reaching the goal.