**(4 points) At least one paragraph per person in the group outlining all experiments you have run. This includes both learning runs and any runs you are doing using heuristics or other methods for comparison. You can only receive full credit if you have some learning results!**

**(4 points) Discuss both the experiments that you have run and your hypothesis on**  
**the outcomes. Analyze the outcome and state any new experiments that need to be**  
**performed and why you think they are important. For example, if you hypothesized**  
**that your learning system was working, and you threw it at the full problem and it**  
**failed, try to come up with a list of possible reasons why it failed and how you could**  
**test these reasons. Note that full credit requires at least one set of learning experiments and results per project partner.**

Alex:

In our exploration of reinforcement learning algorithms applied to the OpenAI Gym Frozen Lake environment, we conducted a temporal difference experiment focused on assessing the effectiveness of heuristics. We proposed that adding heuristics to a baseline Q-Learning algorithm will improve the training rate. Initially, we employed a Q-learning approach, adjusting various hyperparameters such as the learning rate and discount factor to optimize the agent's performance. This experiment shows (See Figure 1) a 90% success rate at around episode 300. To further our understanding and improve performance, we also experimented with heuristic methods, comparing their outcomes against our Q-learning model as a baseline. These heuristics include map boundaries and unneeded traversal to previous states. We found that adding heuristics shows (See Figure 1) a 90% success rate at around episode 220, thus proving our hypothesis is correct for our test cases ran.

**(4 points) Present your results in either graphical or tabular format. Figures: Axes should be labeled, and lines need a legend. All values must show units for that value. Tables should be SMALL in terms of number of cells. Many tables are better served as figures!**

**(4 points) for results that are easily understood, analyzed, and discussed for all members of the project group.**

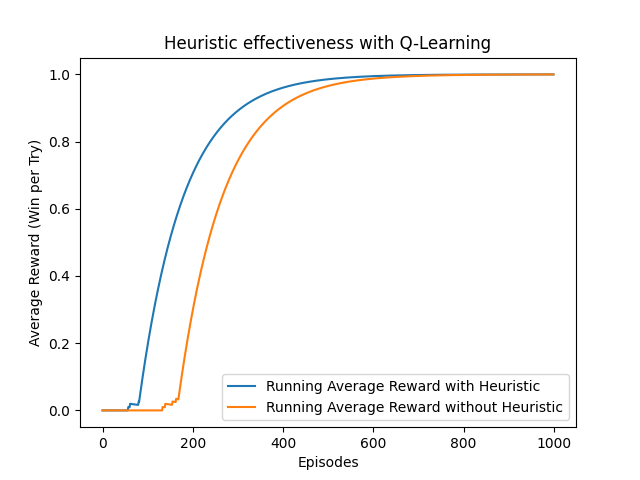


Figure 1: The average reward (number of wins) at each episode for a temporal Q-Learning temporal difference algorithm with heuristics vs. without.

The key results are presented graphically, showing the running average of rewards obtained over episodes. This running average is calculated by averaging the rewards received, with the graph plotting this average as it evolves over the course of training. The graph is expected to show an upward trend, indicating that the agent is learning to navigate the environment more successfully over time. We turned down the number of episodes to 1000 due to the leveling out of both algorithms over time, making the difference hard to see for larger amounts of episodes.

**(1 point) One paragraph per team member discussing any difficulties that you have encountered. Everyone runs into difficulties sometimes and this will enable us to help you. You must say something here to receive credi**t.

Alex:

I am having difficulties finding appropriate material to guide me through understanding how formulas in class connect to actual code. There are not many resources online that fully go into coding these formulas for practical use. I do not think this is a lack of understandable course material, rather it is a lack of in-depth research done by me. Additionally, I am having trouble finding reliable sources that will aid me in the programming process.

Jaiden:

(**1 points) What do you have left to do? I want to see that you can break it down into**  
**manageable step**s.

Alex:

As a group we still need to finish most of the project including but not limited to the final report, comparisons, and the paper reviews. Not only that but we still need to finish two implementations each. Alex still needs to do the eligibility traces along with its respective documentation. In steps, I would say we need to finish our implementations first then start working on the report side of the project.

Jaiden:

**Brevity is the soul of wit. Your total writeup may be no longer than one page per group**  
**member. Longer write-ups will not be grade**d.