Student name: Hien Le
Student Number: 101044264

COMP 4106 - ARTIFICIAL INTELLIGENCE
ASSIGNMENT #3 - REPORT

1. Summary

In this assignment, Learning Automata (LA) and Reinforcement Learning (RL) strategies are implemented in order to solve the Elevator Problem. In other words, LA is used to find an optimal floor for the elevator to park so that it minimizes the waiting time for the next passenger. Waiting time is computed as following:

$$f_i = 0.8 \cdot Q(i) + 0.4 \cdot CEIL[\frac{Q(i)}{2}] + h$$

2. Assumptions and Implementation

Assumptions:

- There are 6 (number of floors) possible actions $\alpha \in \{1, 2, ..., 6\}$
- There are 2 responses {0, 1} (0 for reward, and 1 for penalty).
- h is generated from a Gaussian distribution with standard deviation = 0.5
- Each LA is assumed to be converged after 10,000 iterations of training
- Accuracy is perform on the next 1,000 iterations (from 10,001 to 11,000)
- Speed is calculated as number of iterations needed to converge to 85% accuracy
- All of the measurements including accuracy, speed, and waiting time are calculated by taking the average over an ensemble of 100 experiments

Parameters for each LA schemes:

- Tsetlin Automaton with *memory = 10*
- Krinky Automaton with *memory = 10*
- Krylov Automaton with *memory = 10*
- Linear Reward-Inaction (LRI) with *lambda = 0.2*
-

3. Results and Discussion

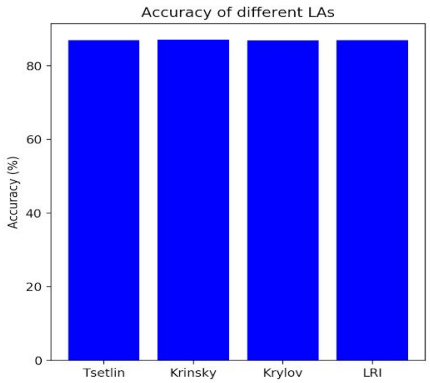


Figure 1. Accuracy

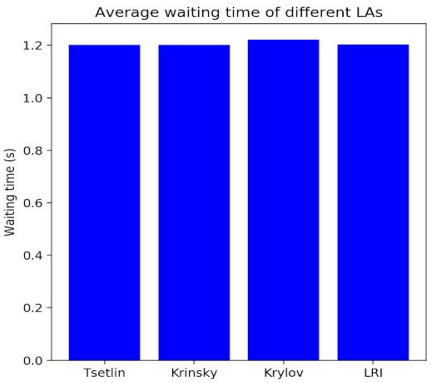


Figure 2. Waiting time

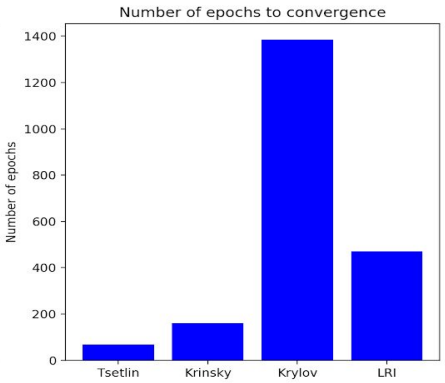


Figure 3. Convergency

The above results were computed after running an ensemble of 100 experiments for each LA.

- The accuracy was approximately 87% (figure 1) for all automatas, thus they all equally work well.
- Average waiting times (figure 2) of each LAs were also very close (~ 1.2). This is an expected behavior because the optimal floor is the one that has the lowest G(i), so substitute G(i) = 1 into the equation:

$$0.8 \times 1 + 0.4 \times CEIL(\frac{1}{2}) + h = 1.2 + h$$

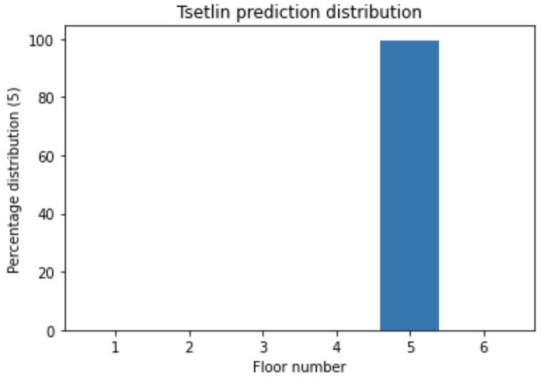
- According to figure 3, Krylov scheme took the longest time (~ 1,385 iterations) to converge while the Tsetlin scheme required the least number of iterations to converge (~ 68 iterations). Therefore, based on their speed, the schemes can be ranked (fastest to slowest) as follow:

$$Tsetlin > Krinsky > LRI > Krylov$$

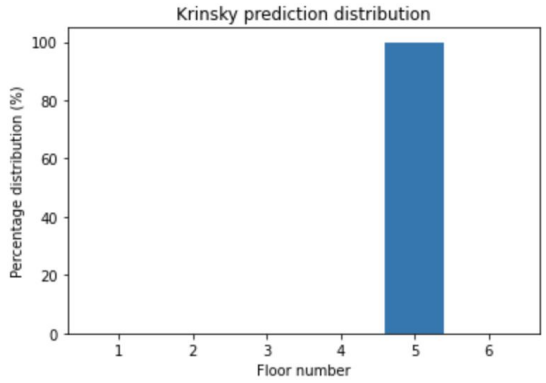
The reason is that the Krylov scheme when receiving an unfavorable response has a 50% chance to turn that response into a reward. As a result, there is a high chance of getting rewarded for doing a wrong action, thus Krylov scheme takes longer to converge. On the other hand, Tsetlin is steadily taking one step at the time toward (or away from if it is a penalty) the correct action and there are only 6 actions in this problem, thus the Tsetlin machine converges faster.

Predictions distribution of each automata:

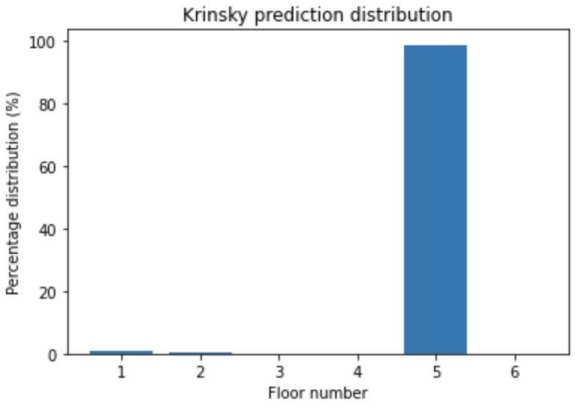
- Tsetlin:
  - `[0.08, 0.09, 0.09, 0.09, 99.54, 0.09]`



- Krinsky
  - `[0.00, 0.00, 0.00, 0.09, 99.9, 0.00]`



- Krylov
  - [0.63, 0.27, 0.09, 0.09, 98.91, 0.00]



  - LRI
    - [0.27, 0.27, 0.36, 0.18, 98.54, 0.36]