Learning Automata

Elementary Reinforcement Learning

Prakriti Kharel
101031299

Environment

The environment returns either a penalty or a reward. For all machines, it decides whether to send a penalty or a reward using the f(i) function as given in the assignment requirement.

If the floor that it came from (the last place it dropped off a passenger) - where the guess was (sent by the LA) < f(i), then it will send a reward, else it will send a penalty

Parameter Used

For Tsetlin, Krinsky and Krylov machines, the state is stored in a 6 * 60 double array. If the state is between 0- 9, then the best elevator to wait is first floor, if the state is between 10-19, then second and so on.

For LRI, the probability of each floor is stored and updated as the simulation runs more instances. If its a penalty, it returns with no changes. If its a reward, then for all the other floors (besides the guess), the probability is decreased by the guess's probability and the probability of guess is increased by the sum of all other probabilities after they have decreased.

Accuracy

For all machines, after 9000 instances, the number of times each floor occurs is stored and the average per experiment is calculated for the next 1000 instance. This is done because after 9000 instances, it usually converges to one to two floors.

To get the final average, the simulation is ran 100 times, and the average of the average is calculated.

Accuracy Order (Most accurate to least accurate)
Krylov, LRI, Tsetlin, Krinsky,

Speed Order (Fastest to slowest)

Tsetlin, Krinsky, LRI, Krylov

Sample Averages for each

Tsetlin Machine (time = 4 seconds average)
The average for one is 0.9541399999999994
The average for two is 0.0
The average for three is 0.0
The average for four is 0.0
The average for five is 0.0
The average for six is 0.04486

Krinsky Machine (time = 5 seconds average)
The average for one is 0.7994399999999999
The average for two is 0.0
The average for three is 0.0
The average for four is 0.0
The average for five is 0.0
The average for six is 0.19956

Krylov Machine (time = 16 seconds average)
The average for one is 0.0
The average for two is 0.0
The average for three is 0.0
The average for four is 0.0
The average for five is 0.0
The average for six is 0.9989999999999989

LRI Machine (time = 13 seconds average)
The average for one is 0.0
The average for two is 0.0
The average for three is 0.0
The average for four is 0.0
The average for five is 0.0
The average for six is 0.9989999999999989