IKT440 – Assignment 1 learning insights

(The Goore Game with Tsetlin Automata)

- How to model an environment that decides on reward and penalty for different actions.
- What kind of random environment we have in the assignment , is it static or dynamic? Why?
- If all voters are independent as given in the assignment, how the learning of best action would be if they were not independent and know each other's votes. (in the Goore game assignment they do not.).
- Why we need to run a random experiment 100 times and take the average?
- What is the difference between an TA interaction steps (voting rounds) which you should perform many times and the 100 runs to take an average of all experiments?
- If we define values for c_1 and c_2, what they really represent? And how that relates to the observed selected action from the TA after it does multiple interaction steps with the environment.
- How we define a randomness in python.
- As you learned in morning lecture. The basic three elements that define a learning problem are: Task, Performance Metric, and the Type of Experience. Please think about where each of these elements fits in the assignment. Especially, what can be your performance metric?
- What kind of learning the TA perform in the Goore game? Active or Passive, and what each kind means?
- Always remember that probabilities of all possible outcomes of an event must sum to 1. Therefore, probability of environment possible responses to each action (an action is as an event) must sum to one. So, c_1 is the probability of penalizing action 1, other possible outcome of action 1 is to receive a reward from the environment. So, probability of action 1being rewarded = (1 – c_1). Same goes for action 2. Hence, do not think that c_1 + c_2 should sum to 1. They are probabilities of two different events (actions). So we can have c_1, c_2 = (.1, .7) and 1 - c_1, 1 - c_2 =(.9,.3).

- Finally, what does it means if we have an environment c_1, c_2 = (.5, .5). Can the TA learns anything from that environment? Why?