## End Course Test | Programming Based Questions

### Upper Confidence Bound Selection Method

**Test Case:** In the upcoming US presidential elections in 2020, Andrew Yang is one of the candidates. You can find a brief information about him and other candidates at <https://www.nytimes.com/interactive/2019/us/politics/2020-presidential-candidates.html>. As usual, Andrew Yang also has a team that helps him run his campaign to gain more popularity. He has proposed a unique idea of *Universal Basic Income* wherein every American will be provided with $1,000 per month as a basic income and this will be funded by the government. People have mixed reviews about this and so his marketing team has come up with an idea of *“Dinner with Yang”*. Since his proposal targets the low income group, he will be hosting dinners across the nation with the target group and will have a discussion with them about their problems as well as collect reviews on his proposal. The marketing team has come up with an idea of promoting this series of dinners on a dedicated website. For this they need your help in choosing the right image to be put on the website – one that will attract public as well has help sympathize with the public which will be leveraged as votes in favor of Yang. They have a list of 20 images which were reviewed by 3 Lakh people / respondents in person as well as on social media. The respondents gave their ratings on a scale of 1 (the image does not look sympathizing) to 5 (the image does looks absolutely sympathizing) for all the 20 images. As an RL expert you are expected to help them maximize the turnout of people on Yang’s dinner from this survey data by rotating different images on their website. The data is in the file named “Andrew\_Yang\_market\_survey\_image\_for\_dinner.csv”.

*Hint: The objective is to help the team maximize reward by rotating different images out of 20 on the website.*

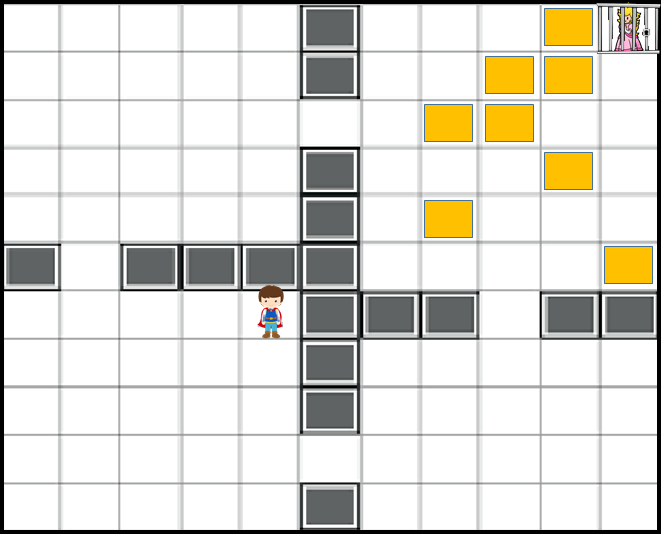
### Epsilon Greedy Selection

**Test Case:** XYZ Company is building an intelligent and relevant search engine focused on the specific needs of lawyers across Europe. They have integrated a recommendation system in the search engine which is rule based as of now. They want an RL algorithm to enable smart suggestions / recommendations to the lawyers based on their profession type e.g. criminal, civil, divorce etc. They have historical data about the number of searches made by the lawyers as per their profession type. They would like your help to identify which profession type they should focus on first as a start. The data is available in “Lawyer\_search\_engine.csv”.

*Hint: You may want to calculate the probabilities first from the frequency count mentioned in column C of the csv file.*

### Policy Evaluation by direct solution of Bellman Equation

**Test Case:** Consider the grid world as shown below. The prince has to find its way from start point to the end point while avoiding the wall (grey squares) as it will give a penalty of -200 and cause the game to restart and the game over area shaded with orange which will lead to a failed status. Can you help the prince find his princess?



### TD Prediction

**Test Case:** The prince has a new challenge this time. He has to collect the key first and then go save the princess. However, he also has a timer! Can you help the prince save his princess in less than 200 iterations?

*Hint: You may want to iterate values of gamma and epsilon for faster learning.*

