# Step 1: To organize the tournament

### Propose a data structure to represent a Player and its Score

A player is characterized by its name/pseudo, its score, and its rank. The name is going to identify and differentiate players. The rank is going to classify them in order to make the groups and eliminate the last ones. The score is mandatory to update the rank. We have 3 attributes so the best data structure to store them in one object player is to create a class Player.

When we initialize a Player, we give him a name, his score is set to 0 and his rank depends on how many players were initialized before him even though at the beginning it doesn’t represent anything.

We actually took the time to create each one of the 100 players of the tournament. We picked the pseudo of the top 100 streamers in the world and wrote them in a text file participants.txt. Then with a dedicated function, we read this file and put every name in a list. After that we looped 100 times the Player class initialization (Player(name, rank=ith iteration)) to create our set of 100 Player objects.

We preferred this class structure to a simple dictionary (key:value)=(name:score) because we wanted to have this third attribute being the rank and also because we felt more comfortable using class. The data persistency of a class object, the possibility to add functions within the class related to the object (Update score, Reset Score, Update rank, etc.) and the ease of use and calling are all arguments that strengthened our choice.

### Propose a most optimized data structures for the tournament (called database in the following questions)

Now that we have our Player structure, we need a tournament structure to organize and classify them through different games. The main purpose of this database is to be able to reach any player with a log complexity. And a simple list of the 100 players is not going to match this constraint because if we are looking for a player that is at the end of this list, then it is going to take longer than a log complexity to reach it.

Since we want to classify players according to their rank, we have an integer value which can be compared to others. In this case, a Binary Search Tree reveals itself to be suited to the situation. When we have to insert data in the tree, at each node (if not the root), the inserted data has two options depending on the value of the rank: if the rank of the inserted player is higher than the rank of the player in the node, then it goes in the right subtree; left otherwise. And if there is no subtree, then the inserted player creates a new one.

With this database, searching for an element becomes easier because we always have only two choices: either go right if the value we are looking for is higher than the one in the actual node, or go left, until we reach the exact same value (if it exists).

However, BST is not the best suited structure. And this is because if by any mean the values we insert in the tree are always higher than the previously inserted value, then we are going to find ourselves with a tree not really looking like a tree because it only has one right subtree at each node. Therefore, if the value we are seeking in the tree is at the end of it (the last right subtree), then our tree will be no better than a classic list and we will reach a linear complexity. In conclusion, we need to balance our tree.

In this case, the AVL tree reveals itself to be the most suited structure for our tournament. Indeed, AVL tree is a balanced Binary Search Tree, meaning the insertion of data works exactly like a BST but it differentiates itself in the balancing. If the right subtree of a node gets 2 subtrees higher than the left, then we apply rotations to rebalance the tree.

With this technic, the AVL tree search complexity cannot exceed log2(n) which is our requirement.

Finally, we have to notice that Nodes of our tree are not just composed by a left subtree, a right subtree, a height (for AVL) and a simple value data. In our case, it is more complex because our data are Players which are structures. The main difference is going to play inside the AVL tree functions : instead of comparing data (players) directly, we will need to compare players.rank.

### Present and argue about a method that randomize player score at each game (between 0 point to 12 points)

In the Player Class, we implemented an UpdateScore method which, with a simple integer randomizer without weighted score, gives a score between 0 and 12. However, to keep the simulation as real as possible, we need to make scores like 10, 11 or 12 less probable than 4,5 or 6. We can use the technique of the dice increasing probability for average score and decreasing for extreme score (0,1,2,11,12). So here we have two “dices” containing numbers between 0 and 6 (included), hence a possible set of score in the range (0,12) (12 included).

Now, it is required that players score is not the sum of all scores from all their games but rather the mean of scores of all their games. Therefore, we cannot just add the randomized score from a game to the previous score, we have to multiply the previous score by (n-1)/n, n being the number of games played, and add to this the new score from this game divided by n. Then we have the mean score.

### Present and argue about a method to update Players score and the database

After updating Players score, we must update the database because the ranking of each player has really strong chances of changing. And to do so, we cannot keep our current AVL tree and change each player’s position in it, this would be overly complicated. If the 100th player scores 12 points, then we have to check if the player above him still has a higher score. If this is not the case, then player 100 exchanges places with player 99 and so on until player 100 finds a player that has a better score than his. And if you think this would be long, we must repeat this process as much times as there are players in the Tree. The conclusion is that it would take forever, and it is not a good solution to update players score within the tree.

However, if we just do an in-order search on the AVL tree, it will give us back the list of all players ordered by their ranking from the worst to the best. Then, we use this list to update players score simply looping iteratively and using UpdateScore function. Finally, with this list we create a new AVL tree which is going to class them upon their score and not their ranking (because it would not be updated at this time). Once all players are inserted inside the tree, another simple in-order search algorithm on the tree would give us the opportunity to rank them: the first one in the list being the 100th of the tournament and the last of the list the first.

It is important to have this subsidiary list alongside the AVL tree because it is this one that we are going to use to update scores and ranking of players.

### Present and argue about a method to create random games based on the database

The 3 first random games are warm up games. Their goal is to put players in the atmosphere of the tournament without any pressure of being kicked out at the end. To prepare them in the best way possible, they should not play again the exact same person 3 times in a row because this would not be representative of the average level: some groups of 10 might be more homogeneous than others, then giving the wrong impression to gamers (either the level is too high or too low). Hence the importance of well mixing the players in those 3 games.

For the first random game, all players score is set to 0 so we can’t know their level. Plus, nobody played nobody so far so we can create games of 10 players however we want, there would not be redundancy. Then it is sold, players will play against each other depending on their ranking or more precisely the alphabetical order because this is how their first ranking was based on. And to get players in a ranking order from the AVL Tree, we have the in-order search algorithm.

Now for the second random game, scores are updated, so are ranks and the new AVL Tree is built. We want an algorithm that is going to give us back the list of players from this tree but in a different order from the previous one. And even though doing an in-order search algorithm would not give us the same list (in the same order) as earlier, this would not respect our wish of heterogeneity among the random games. To keep the splitting as random as possible we will use the fact that the ranking changes of course but also and mainly use a different search algorithm: pre-order or post-order. Here we use the pre-order search algorithm to organize the second random game.

For the last game, we repeat the same steps as earlier but apply the post-order search algorithm instead.

Even though we didn’t really need scores and ranking for the creation of the 3 random games, we still needed to update them because they are mandatory for the rest of the tournament. Indeed, at the end of the 3 random games, the 100 players will be ranked according to their score so that the tournament and the disqualification games can begin. (the score of each player is not reset after the 3 games, it wouldn’t be fair for top players).

### Present and argue about a method to create games based on ranking

For the creation and repartition of players according to their rank, we had to make a choice:

* Rewarding the best players by making them play against the worst players
* Giving all players, good or bad, the same chance of success batching the best players together and the worst players together

The second option is simple: we get the list of players in-order from the AVL tree then split the list every ten players so that player ranked 1 play against 2, 3, 4 until 10, and players 91 to 100 play against. Unfortunately, this technic does not reward the best players, it evens disadvantages them. The 10th player for example will be the worst player from his group and will not get the chance to score as many points as he could have if he had played against worse players.

However, since we are giving points randomly to each player and we do not respect some type of structured pointing system, this logic of rewarding gets blown away. Therefore, we decided to go for the simplest way of splitting groups from 10 to 10 rank-wisely. The process is quite straight forward: get the list of rank-ordered players with the in-order algorithm, go through the list with a for loop and place players in a group (list), changing list every 10 players.

### Present and argue about a method to drop the players and to play game until the last 10 players

Serious things begin. Shuffle games are over, and the official tournament takes place: 9 rounds of one game after each the 10 last players (least score) are permanently eliminated. 9 rounds because we need to eliminate 90 players to keep the 10 best for the final step of the tournament.

The scheme is the following: for the first round, we have our AVL Tree and our affiliated rank-ordered list.

* we update the scores (with the list) after the game
* we create the new AVL Tree
* we get the list of players ordered from the lowest score to the highest
* we update the ranks
* we delete the 10 first players of the list (10 worst)

So, we repeat this action 9 times.

The reason we do not practice a deletion directly on the AVL Tree is because when we create the AVL tree after the score update, we do not know the new rank of players. So, we cannot go through the AVL tree searching for the 10 worst players according to their ranking because it will not be up to date and we could find ourselves to eliminate the 89th player because his rank was 91 before the game. Furthermore, we cannot do the same thing with the score because we cannot know the exact score for which the 10 last players are below.

Hence the need to get the new list of ordered players (in-order algorithm) from the AVL Tree to update the ranks and delete the 10 worst players from it. At this point, we could question the utility of the AVL Tree since we don’t practice any action on it like deletion. But in fact, the AVL Tree is super effective for classifying players and getting the list of players in order quickly.

### Present and argue about a method which display the TOP10 players and the podium after the final game.

When only the 10 best players are remaining, the tournament organization changes. All scores are reset to 0 to not disadvantage the 10th player who has a lower score than the 1st. Players are going to fight for the podium around five games. After each game, scores are always updated, but ranks and also the AVL tree don’t need to be because we don’t attach any importance to the classification of players until the fifth and last game is played since there are no disqualifications.

Therefore, after the final game, with the list of 10 players, we create our AVL Tree inserting players according to their score and not to their not updated rank. Then we get the same list but ordered from the worst to the best with the in-order algorithm. Finally, we loop through the list starting from the end and display the rank, score and name of players from the best (1st) to the worst (10th).

# Step 2: Professor Layton < Guybrush Threepwood < Us

### Represent the relation (have seen) between players as a graph, argue about your model.

We are given several information:

* When a player sees another player, they always both see each other, there can’t be one player seeing another without being seeing. This has for consequence that the adjacency matrix will be symmetrical.
* Impostors can’t see each other

To register this information on who saw who, we decided to create an adjacency matrix after each crewmate death. This matrix is squared and has as many rows and columns as living players in the game (we count in the player who just got killed). This matrix is a Boolean matrix saying if a player saw (True) or not (False) another one (1 if seen; 0 if not seen).

For example, if we have 10 players and one got killed, we have a 10\*10 matrix of Boolean. If player 0 who died has seen the players 1, 4 and 5, then the value of the matrix is set to 1, same for and .

We also have the , and values of the matrix set to 1 but it is not necessary to write them since the matrix is symmetrical. However, in the python code, it will be easier to complete the whole matrix because when we will count who saw who, it will be easier to count on one column or one line instead of one part on one line n and the other on column n.

Here is the representation of the Adjacency Matrix of the example given in the subject.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0 |  | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 1 |  |  | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 2 |  |  |  | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| 3 |  |  |  |  | 1 | 0 | 0 | 0 | 1 | 0 |
| 4 |  |  |  |  |  | 0 | 0 | 0 | 0 | 1 |
| 5 |  |  |  |  |  |  | 0 | 1 | 1 | 0 |
| 6 |  |  |  |  |  |  |  | 0 | 1 | 1 |
| 7 |  |  |  |  |  |  |  |  | 0 | 1 |
| 8 |  |  |  |  |  |  |  |  |  | 0 |
| 9 |  |  |  |  |  |  |  |  |  |  |

### Thanks to a **graph theory problem**, present how to find a set of probable impostors.

Logically, if a crewmate dies, then the impostor has to be one of the members he has previously seen. Therefore, if 0 is found dead, then the impostor is either 1, 4 or 5, or maybe two of them. With this probable set of impostors we can estimate who is going to be the second impostor. Indeed, if for example, 1 is the impostor, and we know impostors don’t see each other at all, the set of probable impostors is all the players 1 didn’t see : 3, 4, 5, 7, 8 and 9. In our case, it isn’t really helpful, but in a common game players ought to see more than 3 other players increasing our chances of finding the impostors with this technic.

Furthermore, we can’t make assumptions after the first kill. A better and more efficient way would be to store this information and then add it to the new information we get after the second kill. To do so, we can create two new attributes for players:

* The first one is a coefficient of Impostorness that increases of a certain amount depending on the number of other players the dead one saw. To take back our example, if 0 dies, then 1, 4 and 5 see their coefficient increased by 1 divided by 3. After the second kill, if 9 dies and he saw 1 and 4 then those players see their coefficient increased by 0.5. At this point, 1 and 4 both have coefficient equal to 0.83. They can both be impostors or only one of them.
* The second one is a list containing the list of players seen at each round. This is useful because considering one dead player saw only two players and one of them died later in the game, then the other must be the imposter. Plus, if one impostor is found, then we can exonerate all players this impostor saw since impostors don’t see each other. Hence the need to store this information and to update it after each kill.

Following these steps, if we arrive at a point where a suspected player has seen every other player, except one, then we can highly deduce that this player and the one he has never seen are the impostors. Reciprocally, if a player is a suspect with a high Impostorness coefficient but has seen every other player alive, then he cannot be an impostor.

This technic combined with the Impostorness coefficient can get us to find impostors way quicker.

Let’s get back to our example. 0 was killed and he saw 1, 4 and 5:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | IC | sum |
| 0 |  | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 |  |  | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1/3 | 1/3 |
| 2 |  |  |  | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 3 |  |  |  |  | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 4 |  |  |  |  |  | 0 | 0 | 0 | 0 | 1 | 1/3 | 1/3 |
| 5 |  |  |  |  |  |  | 0 | 1 | 1 | 0 | 1/3 | 1/3 |
| 6 |  |  |  |  |  |  |  | 0 | 1 | 1 | 0 | 0 |
| 7 |  |  |  |  |  |  |  |  | 0 | 1 | 0 | 0 |
| 8 |  |  |  |  |  |  |  |  |  | 0 | 0 | 0 |
| 9 |  |  |  |  |  |  |  |  |  |  | 0 | 0 |

After the first kill, we cannot conclude on who are the impostors.

4 was killed and he saw 1 and 7.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | IC | Sum |
| 0 |  |  |  |  |  |  |  |  |  |  |  |  |
| 1 |  |  | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1/2 | 5/6 |
| 2 |  |  |  | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 |
| 3 |  |  |  |  | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| 4 |  |  |  |  |  | 0 | 0 | 1 | 0 | 1 | 0 | 1/3 |
| 5 |  |  |  |  |  |  | 0 | 0 | 1 | 0 | 0 | 1/3 |
| 6 |  |  |  |  |  |  |  | 0 | 1 | 1 | 0 | 0 |
| 7 |  |  |  |  |  |  |  |  | 1 | 1 | 1/2 | 1/2 |
| 8 |  |  |  |  |  |  |  |  |  | 0 | 0 | 0 |
| 9 |  |  |  |  |  |  |  |  |  |  | 0 | 0 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0 |  | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 1 |  |  | 2 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| 2 |  |  |  | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| 3 |  |  |  |  | 1 | 0 | 1 | 0 | 1 | 0 |
| 4 |  |  |  |  |  | 0 | 0 | 1 | 1 | 1 |
| 5 |  |  |  |  |  |  | 0 | 0 | 1 | 0 |
| 6 |  |  |  |  |  |  |  | 0 | 2 | 1 |
| 7 |  |  |  |  |  |  |  |  | 1 | 1 |
| 8 |  |  |  |  |  |  |  |  |  | 0 |
| 9 |  |  |  |  |  |  |  |  |  |  |

Here, the dead crewmate only saw 2 other players, so the probability one of them is the impostor is high (0.5). We add those probabilities to the sum of coefficients and get 5/6 for 1 and ½ for 7 meaning 1 has more chances to be the impostor. Yet, we can’t deduce with certainty that 1 is the impostor. But if we could prove 7 is innocent, then we would be sure 1 is an impostor. And one way to do this is to verify if 7 has seen all other players. If that is the case, then he can’t be an impostor since two impostors cannot see each other. This technic only works when two impostors are remaining.

Furthermore, when an impostor is unmasked, then we have to update all previously calculated Impostorness coefficient. Indeed, if 1 is unmasked, then we are sure all players he saw during the game are not impostors. If those exonerated had an Impostorness coefficient different from 0, then we have to redistribute their value to those not innocent.

If we look at the matrix of the sum of meetings, we see that 8 saw every one except 9 and 1 who is a serious suspect. Therefore if 1 is proven to be an impostor then its Impostorness coefficient would be transferred and split between all players he didn’t see and if it remains only one player then he has to be the impostor.

### Argue about an algorithm solving your problem.

To make sure our investigation technics work, we wanted to try it on a real situation. However, a real situation would be to play actual games of Among Us, be able to detect when players see each other and mostly that they respect all the assumptions we made (for example, the fact impostors can’t see each other isn’t representative of the reality because they can sometimes meet up).

The plan is that we choose the 2 impostors before the program launches so we know who they are but the computer doesn’t and his goal is to unmask them.

So, we randomly choose 2 impostors in the set of 10 players and then repeat those steps until no impostors are remaining or the number of impostors equals the number of crewmates:

1. create Adjacency matrices:

* We randomly set values of 0 and 1 in the matrix
* We make sure the matrix is symmetrical
* We also are careful to not putting 1 between 2 impostors (meaning they saw each other)

1. Register that information in player objects:

* IDs of seen players are put in a list
* this list is appended to the list

1. randomly kill one crewmate (any alive player other than an impostor)
2. Update Impostorness coefficients for this round and the previous ones
3. Reveal an impostor if there are enough evidence to do so:

* Either we are 100% sure one player is an impostor so we unmask him
* Or we are at a point where if the argument doesn’t unmask an impostor, the game ends (for example 3 crewmates and 2 impostors remain, if no impostor gets discovered, then next round there will be one crewmate left making it impossible for cremates to win.

List\_players\_alive=[0,1,2,3,4,5,6,7,8,9]

Impostors=[1,7]

While (not game\_ended)

New\_matrix= Get\_Matrix\_meetings(list\_players\_alive)

We get the list of players the killed player saw

List\_Suspects=Suspects(New\_Matrix)

If this list contains only one player, then he has to be an impostor

If (lengthList\_Suspects == 1)

Impostor=List\_Suspects[0]

Busted\_Impostor+=1

We update Impostorness coefficients

Update\_Impostorness\_Coeff(New\_Matrix\_With\_Dead)

Sum\_matrix

### Implement the algorithm and show a solution.

Implement the previous algorithm

# Step 3: I don't see him, but I can give proofs he vents!

### Presents and argue about the two models of the map.

**First model : without vents**

Undirected graph : players can go in each rooms in both ways (from B to A and A to B), each way, aka edges, between rooms, aka nodes, is weighted by the amount of time needed to go from one room to another.

To calculate weights, we need to measure the distance between rooms connected to each other.

To estimate who is the impostor, we need to calculate the time a player put to go from one room to another. If this time is less than the weighted edge, then the player is an impostor.

**Second model : with vents**

This graph is a subgraph of the previous one, meaning we are going to take the same one and add edges to it. And the edges we are adding are the vents travels impostors can do between two rooms. Because the time to travel in vents is null, the weights on these new edges are all 0.

### Argue about a **pathfinding algorithm** to implement.

Matrix of size m\*m with m the number of rooms. The mij element of the matrix represents the minimum amount of time to travel from room i to room j.

### Implement the method and show the time to travel for any pair of rooms for both models.

Create a function that takes in parameter a depart room and an arrival room and returns the time of travel with the Dijkstra algorithm (finding the shortest paths between nodes in a graph. (no negative weights, directed or undirected graph, finite number of vertices, defined source))

### Display the interval of time for each pair of room where the traveler is an impostor.

Function using the previous function and calculate the eta for map 1 and 2 and keep the pairs of room where the eta is different.

# Step 4: Secure the last tasks

### Presents and argue about the model of the map.

The map is going to be a subgraph of the precedent graph. Indeed, at the end of the game, only few tasks are remaining and rooms are no longer needed because no tasks are to be accomplished in those rooms. Therefore, we need to delete nodes (corresponding to rooms with no tasks), and create new edges weighted with the sum of the previous edges.

### Thanks to a **graph theory problem**, present how to find a route passing through each room only one time.

We are going to use an algorithm of the commerce traveler. Finding the shortest or fastest way/itinerary between multiple tasks remaining.

To do so, we use the Hamiltonian path (cf CMO3) : undirected graph, each place must be visited once and only once (visit every vertex/node exactly once)

### Argue about an algorithm solving your problem.

Explain functioning of Hamiltonian algorithm

### Implement the algorithm and show a solution.

Implement previous algorithm with python.