Homework # 10
MATH 110

In this homework, we will work with the attached Enron dataset. You can read the history of Enron here: `https://en.wikipedia.org/wiki/Enron`. The dataset contains the emails of 88 employees. For each employee, the `sent/` directory contains all the emails sent by the employee. Look through the directories to see the directory structure and the email formats.

Our goal in this homework will be to identify the number of emails sent between the different employees.

1. Look at the format of email files in the dataset. The emails all have a similar format to identify sender, recipient, subject line, etc. Create a `Email` object with the following constructor and methods.
   (a) `Email(filename)`, where `filename` is the name of the file in which the email is stored.
   (b) `to.Email(em)` should return the name of the recipient of the emails (found in the "X-To" line of the email).
   (c) `from.Email(em)` should return the name of the sender of the email (found in the "X-From" line of the email).
2. Create a `Employee` object with the following constructor and methods.
   (a) The constructor should have the form `Employee(directory_name)` where `directory_name` is the name of one of the directories found in `enron_dataset/`. For example, "arnold-j" is a particular value that could be used for `directory_name`.
   (b) `get_number_of_emails.Employee(e)` should return the number of emails the employee sent, where `e` is the Employee object.
   (c) `get_email_filename.Employee(e, i)` should return the filename for the `i$th email` sent by the employee.
   (d) `get_name.Employee(e)` should return the name of the employee. For example if the object was formed using `directory_name` equal to "arnold-j" then the employee name is "John Arnold". (Hint: the name of the employee is found in the X-From line of the employees sent emails.)

3. Pick 5 employees. Use the Email and Employee objects and methods to do the following. Construct a matrix that specifies the number of emails sent by one employee to another employee across all pairs of the 5 employees. Use the package `igraph` to convert the matrix into a graph and then plot the graph. (Hint: Write a function that, given two employee objects, determines the number of emails sent by one employee to the otheremployee. Then, write a double for loop - it's ok to do so in this case - to fill the matrix describing all pairs of the 5 employees.)