Assignment 2

By Aditya Anil,AM.EN.U4AIE19006

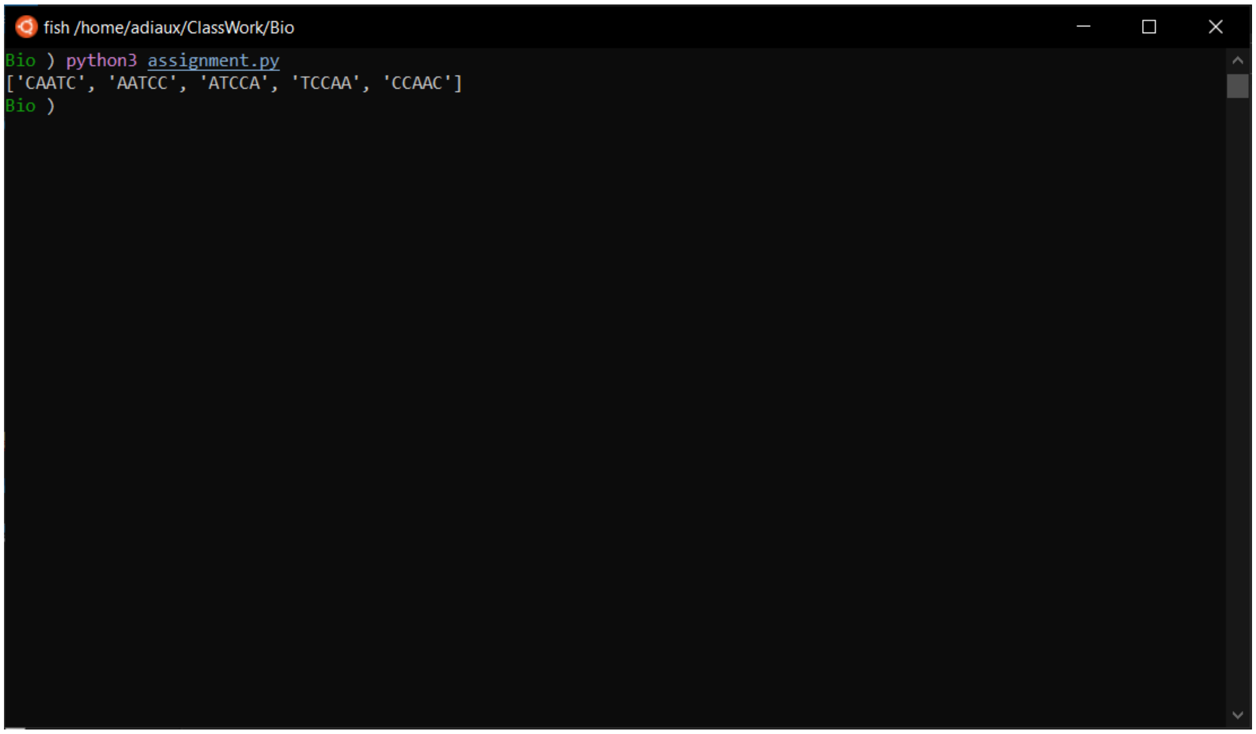
Question 1

String Composition Problem : Generate the k-mer composition of a string.
Given: An integer k and a string Text.
Return: Compositionk(Text) (the k-mers can be provided in any order).

Code

```python
def q1(k,string):
    Kmer=[]
    for i in range(len(string)-k+1):
        Kmer.append(string[i:i+k])
    return Kmer
print(q1(5,"CAATCCAAC"))
```

Output Screenshot



```
fish /home/adiaux/ClassWork/Bio
Bio ) python3 assignment.py
['CAATC', 'AATCC', 'ATCCA', 'TCCAA', 'CCAAC']
Bio )
```

Question 2

String Spelled by a Genome Path Problem: Find the string spelled by a genome path.
Given: A sequence of k-mers Pattern1, ... , Patternn such that the last k - 1 symbols of Patterni are equal to the first k - 1 symbols of Patterni+1 for i from 1 to n-1.
Return: A string Text of length k+n-1 where the i-th k-mer in Text is equal to Patterni for all i.

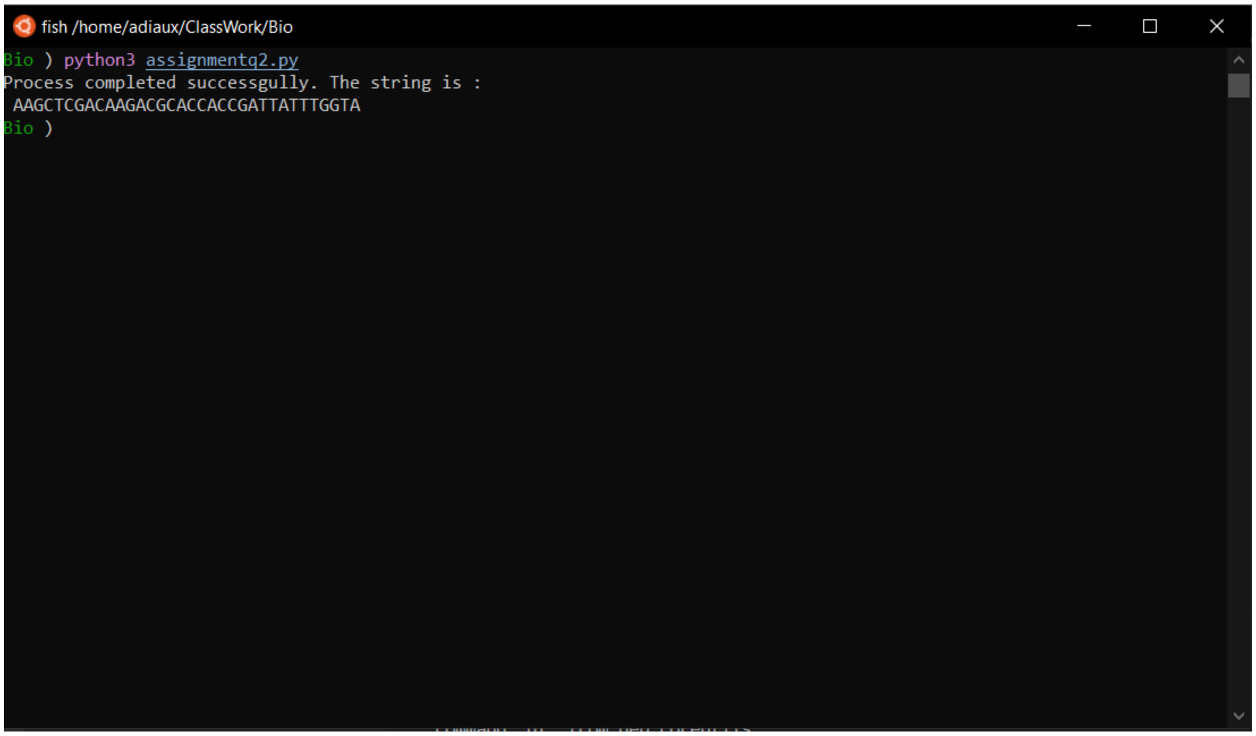
Code

```
flist=[]
def q2():
    file=open('test.txt', 'r')
    first_line=file.readline().strip()
    file.close()
    linList=open('test.txt', 'r').readlines()
    listFinal=[]
    for line in linList :
        listFinal.append(line.rstrip())
    finalString=""
    finalString=finalString+first_line
    del listFinal[0]
    for line in listFinal:
        if(len(first_line)!=len(line)):
            flag=1
        finalString=finalString+line[len(first_line)-1]
        flag=0
    flist.append(finalString)
    flist.append(flag)
    return flist

tempList=q2()
if tempList[1]==1:
    print("The data in file is inconsistent" )
else:
    print("Process completed successgully. The string is : \n",tempList[0])
```

Text file reading data

AAGCTCGACAAGACGCACCACCGAT
AGCTCGACAAGACGCACCACCGATT
GCTCGACAAGACGCACCACCGATTA
CTCGACAAGACGCACCACCGATTAT
TCGACAAGACGCACCACCGATTATT
CGACAAGACGCACCACCGATTATTT
GACAAGACGCACCACCGATTATTTG
ACAAGACGCACCACCGATTATTTGG
CAAGACGCACCACCGATTATTTGGT
AAGACGCACCACCGATTATTTGGTA

Output Screenshot



fish /home/adiaux/ClassWork/Bio

```
Bio ) python3 assignmentq2.py
Process completed successgully. The string is :
 AAGCTCGACAAGACGCACCACCGATTATTTGGTA
Bio )
```

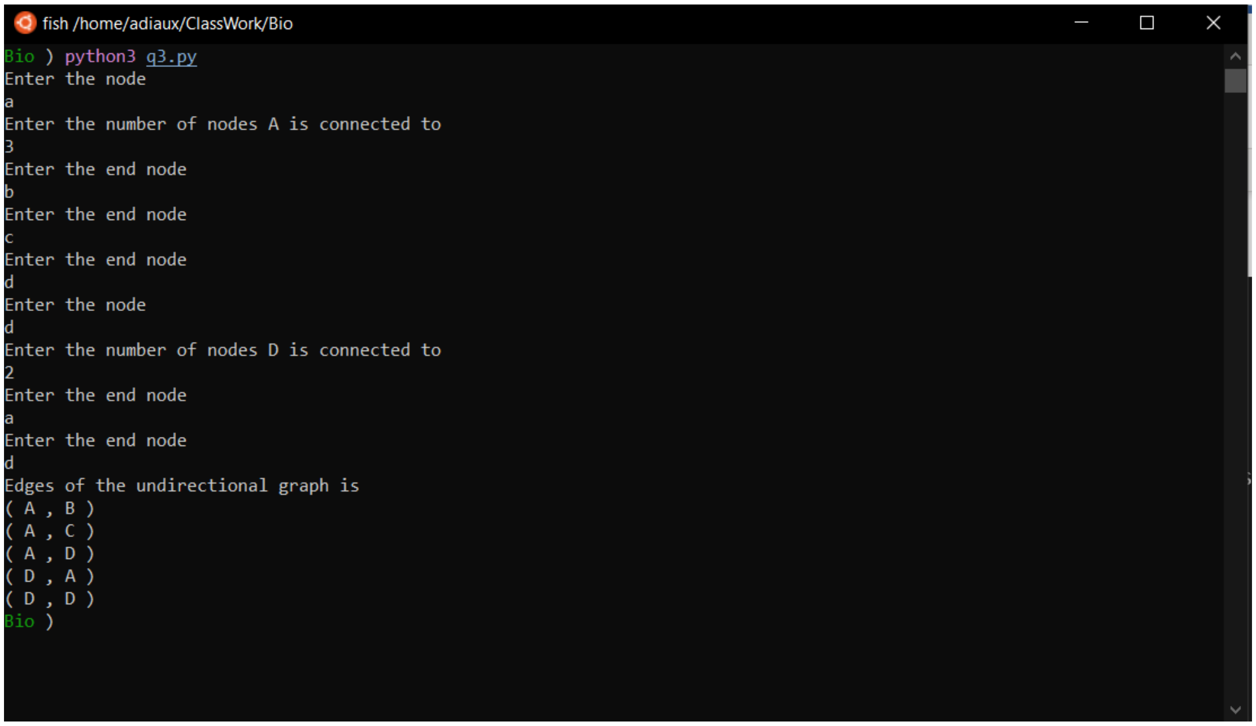
Question 3

Implement graph in python using dictionary data structure in python.

Code

```python
from collections import defaultdict
def defError():
    return "Not Present"

def getInput():
    nodeStart=input("Enter the node \n").upper()
    list=[]
    nodeEndCount=int(input("Enter the number of nodes "+nodeStart+" is connected
to\n"))
    for i in range(0,nodeEndCount):
        list.append(input("Enter the end node\n").upper())

    baseDict[nodeStart]=list
def show():
    print("Edges of the undirectional graph is")
    for key in baseDict.keys():
        list=baseDict[key]
        for i in range(len(list)):
            print("(",key,",",list[i],")")

baseDict = defaultdict(defError)

getInput()
getInput()
show()
```

Output Screenshot



fish /home/adiaux/ClassWork/Bio

```
Bio ) python3 q3.py
Enter the node
a
Enter the number of nodes A is connected to
3
Enter the end node
b
Enter the end node
c
Enter the end node
d
Enter the node
d
Enter the number of nodes D is connected to
2
Enter the end node
a
Enter the end node
d
Edges of the undirectional graph is
( A , B )
( A , C )
( A , D )
( D , A )
( D , D )
Bio )
```