الجامعة الإسلامية العالمية ماليزيا
INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA
يونيورسيتي إسلام انتارابغسا ملسيا

(Company No. 101067-P)

Garden of Knowledge and Virtue

SEMESTER 1, 2021/2022

CSCI 4342 NLP

NATURAL LANGUAGE PROCESSING

Assignment 1

PREPARED BY:

NAME	MATRIC NO.
MOHAMED MOUBARAK MOHAMED MISBAHOU MKOUBOI	1820705
ISLAM MD SHARIFUL	1720601
ABDO ABDELLA MAME	1714883
NUR HANANI BINTI MOHAMAD PAUZI	1916806

LECTURER:

Dr. SURIANI BT. SULAIMAN

Code:

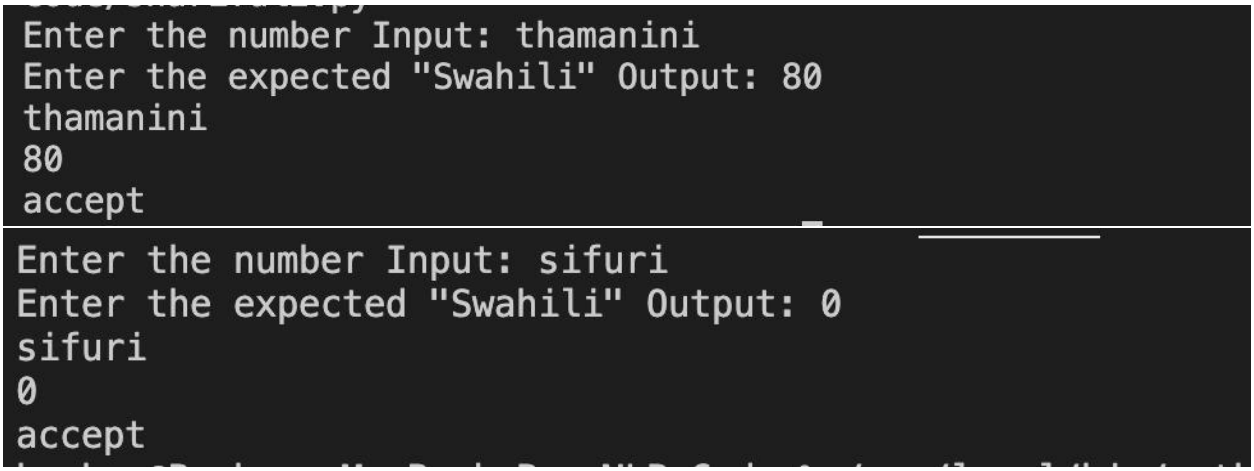
```
"""

@author: MOHAMED MOUBARAK MOHAMED MISBAHOU MKOUBOI    1820705
        ISLAM MD SHARIFUL         1720601
"""
from nltk.nltk_contrib.fst.fst import *
import os
os.chdir('C:\Users\SAMSUNG\Desktop')



class myFST(FST):
    def recognize(self, iput, oput):

        self.inp = list(iput)
        self.outp = list(oput)

        if list(oput) == f.transduce(list(iput)):
            return True
        else:
            # print(outp)
            return False



f = myFST('swahili')

# declare the states
for i in range(1, 5):
    f.add_state(str(i))
```

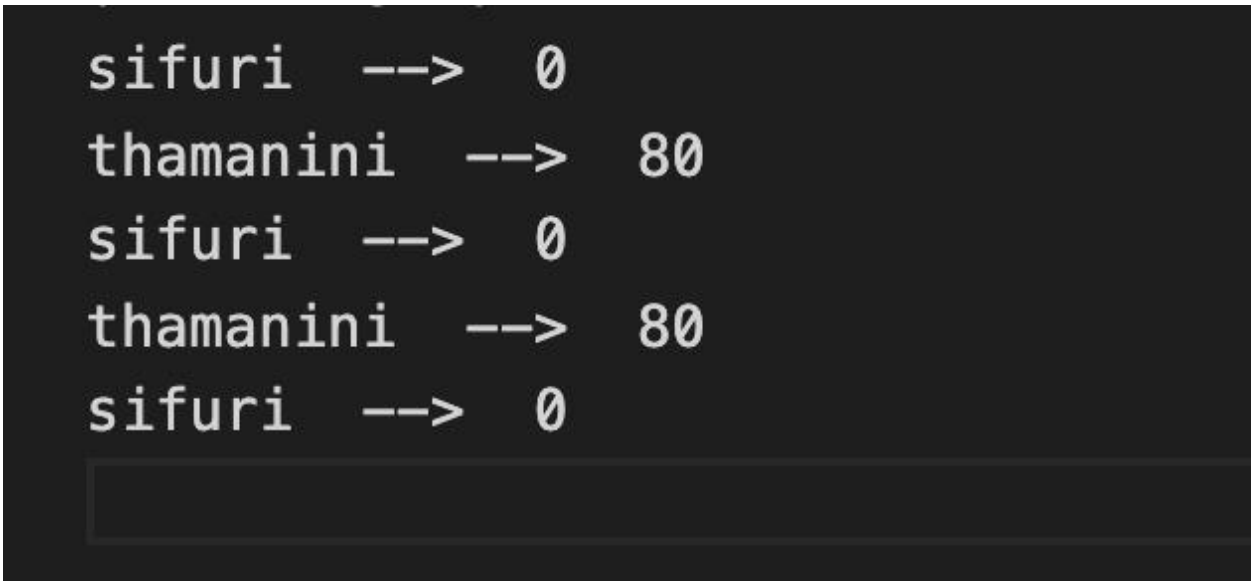
```
f.initial_state = '1'
#f.initial_state = '2'
f.set_final('2')
f.set_final('4')

# setting up the arcs
f.add_arc('1', '2', ('sifuri'), ('0'))
f.add_arc('1', '2', ('moja'), ('1'))
f.add_arc('1', '2', ('mbili'), ('2'))
f.add_arc('1', '2',  ('tatu'), ('3'))
f.add_arc('1', '2',  ('nne'), ('4'))
f.add_arc('1', '2', ('tano'), ('5'))
f.add_arc('1', '2',  ('sita'), ('6'))
f.add_arc('1', '2',  ('saba'), ('7'))
f.add_arc('1', '2', ('nane'), ('8'))
f.add_arc('1', '2', ('tisa'), ('9'))
f.add_arc('1', '2',  ('kumi'), ('10'))
f.add_arc('1', '2',  ('kuumi na moja'), ('11'))
f.add_arc('1', '2', ('kumi na mbili'), ('12'))
f.add_arc('1', '2',  ('kumi na saba'), ('17'))
f.add_arc('1', '2', ('ishrini'), ('20'))
f.add_arc('1', '2',  ('ishrini na tano'), ('25'))
f.add_arc('1', '2',  ('thalathini'), ('30'))
f.add_arc('1', '2',  ('arubaini'), ('40'))
f.add_arc('1', '2', ('hamsini'), ('50'))
f.add_arc('1', '2',  ('hamsini na tano'), ('55'))
f.add_arc('1', '2',  ('sitini'), ('60'))
f.add_arc('1', '2', ('sabini'), ('70'))
f.add_arc('1', '2',  ('thamanini'), ('80'))
f.add_arc('1', '2',  ('tisini'), ('90'))
f.add_arc('1', '2',  ('mia moja'), ('100'))
f.add_arc('1', '2',  ('mia tatu'), ('300'))
```

```
f.add_arc('1', '2', ('mia moja thalathini na sita'), ('136'))
f.add_arc('1', '2',  ('mia tisa tisini na tisa'), ('999'))
f.add_arc('1', '2',  ('elfu moja'), ('1000'))
f.add_arc('1', '2',  ('elfu moja mia tisa tisini na saba'), ('1997'))
f.add_arc('1', '2', ('elfu mbili'), ('2000'))
f.add_arc('1', '2',  ('elfu tano mia nne tisini na nane'), ('5498'))
f.add_arc('1', '2',  ('elfu kum'), ('10000'))
f.add_arc('1', '2', ('elfu mia moja/laki'), ('100000'))
f.add_arc('1', '2',  ('nusu'), ('1/2'))
f.add_arc('1', '2', ('mbili na nusu'), ('2 1/2'))
f.add_arc('1', '2',  ('robo'), ('1/4'))
f.add_arc('1', '2',  ('arubaini na saba na robo tatu'), ('47 3/4'))

# function to check the arcs for availability and writing to file



def translator(inp, outp):
    arcs_file = open('Swahili-Trans.dat', 'a')
    arcs = ""
    arcs += ''.join(inp) + "  -->  "
    if f.recognize(inp, outp):
        print(outp)
        print("accept")
        arcs += ''.join(outp) + '\n'
    else:
        print("reject")
        arcs += ''.join('reject') + '\n'
    arcs_file.write(arcs)



inp = input('Enter the number Input: ')
outp = input('Enter the expected "Swahili" Output: ')
```

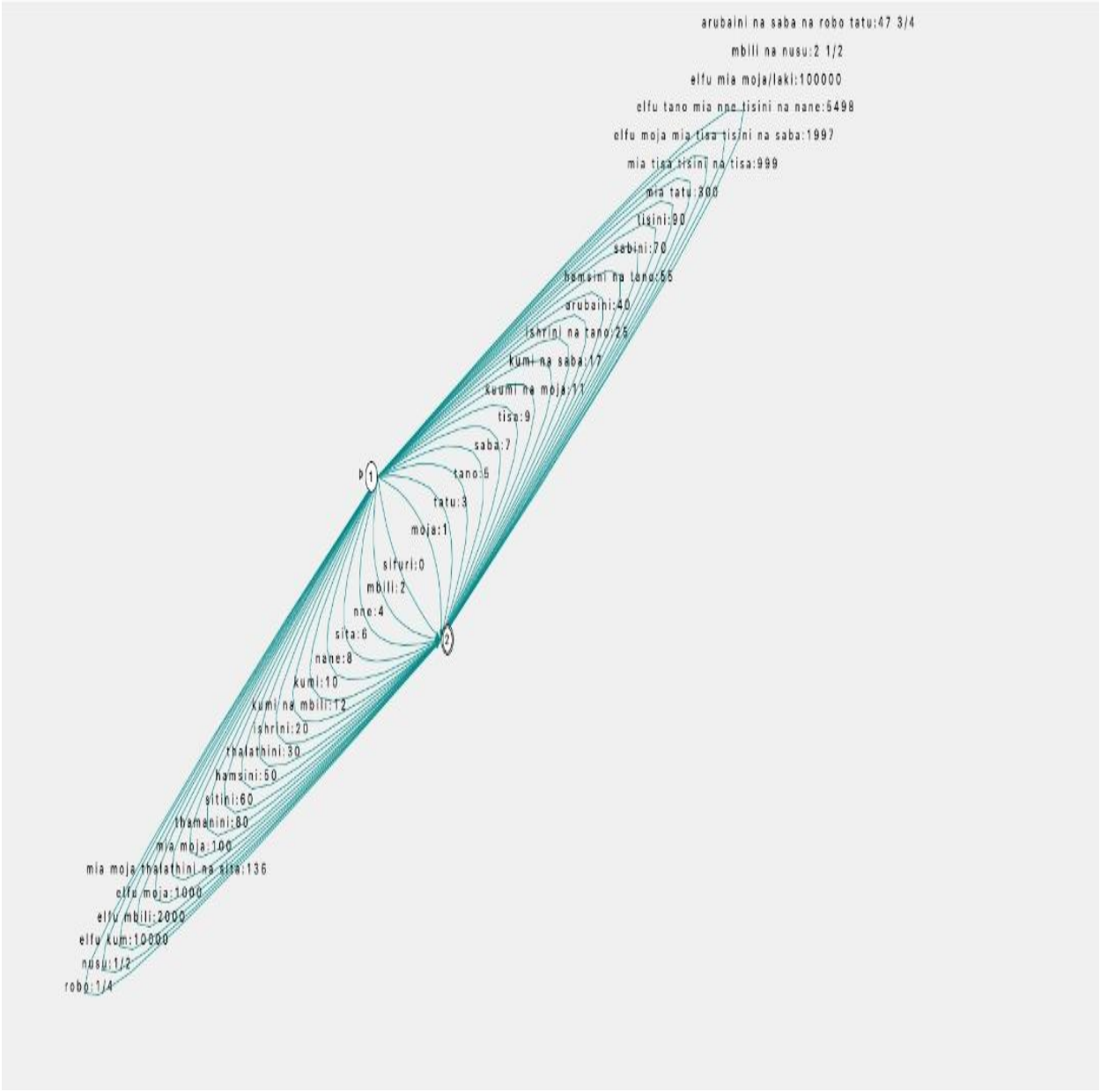
```
print(inp)


# calling the function
translator(inp, outp)


# displaying the fst structure
disp = FSTDisplay(f)
```

Output Snap:





FST construction snap:


arubaini na saba na robo tatu:47 3/4
mbili na nusu:2 1/2
elfu mia moja/laki:100000
elfu tano mia nne tisini na nane:5498
elfu moja mia tisa tisini na saba:1997
mia tisa tisini na tisa:999
mia tatu:300
tisini:90
sabini:70
hamsini na tano:55
arubaini:40
ishirini na tano:25
kumi na saba:17
kumi na moja:11
tisa:9
saba:7
tano:5
tatu:3
moja:1
sifuri:0
mbili:2
nne:4
sita:6
nane:8
kumi:10
kumi na mbili:12
ishirini:20
thalathini:30
hamsini:50
sitini:60
thamanini:80
mia moja:100
mia moja thalathini na sita:136
elfu moja:1000
elfu mbili:2000
elfu kumi:10000
nusu:1/2
robo:1/4
1
2