```
In [ ]:
```

Name: Harsh Zanwar Roll No.: 74 Branch:CSE(AIML)

Aim:4(a) Write a Python NLTK program to perform Stemming and Lemmatization on set of tokens. PorterStemmer(), SnowballStemmer(), LancasterStemmer(), RegexpStemmer(), WordNetLemmatizer(), lemmatizer.lemmatize()

4(b) Write a program to perform Morphological generation on different surface forms of a word

In [15]:
```
import nltk
from nltk.stem import PorterStemmer, SnowballStemmer, LancasterStemmer, RegexpS
from nltk.stem.wordnet import WordNetLemmatizer
from nltk.tokenize import word_tokenize

# Sample sets of tokens with different POS tags
tokens = [("playing", "VBG"), ("played", "VBD"), ("plays", "VBZ"), ("player", '
          ("jumps", "VBZ"), ("jumping", "VBG"), ("jumped", "VBD"), ("jumper", '
          ("running", "VBG"), ("ran", "VBD"), ("runs", "VBZ"), ("runner", "NN")
          ("eating", "VBG"), ("ate", "VBD"), ("eats", "VBZ"), ("eater", "NN"),
          ("talking", "VBG"), ("talked", "VBD"), ("talks", "VBZ"), ("talker", '
          ("singing", "VBG"), ("sang", "VBD"), ("sings", "VBZ"), ("singer", "NN
          ("writing", "VBG"), ("wrote", "VBD"), ("writes", "VBZ"), ("writer", '
          ("swimming", "VBG"), ("swam", "VBD"), ("swims", "VBZ"), ("swimmer", '

# Initialize stemmers and lemmatizer
porter_stemmer = PorterStemmer()
snowball_stemmer = SnowballStemmer("english")
lancaster_stemmer = LancasterStemmer()
regexp_stemmer = RegexpStemmer("ing$|ed$", min=4)
wordnet_lemmatizer = WordNetLemmatizer()

# Perform stemming and lemmatization on the tokens using different stemmers anc
print("Token\t\tPOS\tPorter Stemmer\tSnowball Stemmer\tLancaster Stemmer\tRegex
for token in tokens:
    word = token[0]
    pos = token[1]
    porter = porter_stemmer.stem(word)
    snowball = snowball_stemmer.stem(word)
    lancaster = lancaster_stemmer.stem(word)
    regexp = regexp_stemmer.stem(word)
    lemma = wordnet_lemmatizer.lemmatize(word, pos=pos.lower()[0])
    print(f"{word}\t\t{pos}\t{porter}\t\t{snowball}\t\t\t\t{lancaster}\t\t{rege
Token           POS     Porter Stemmer  Snowball Stemmer        Lancaster Ste
mmer    Regexp Stemmer  WordNet Lemmatizer
playing         VBG     play            play                            play
play            play
played          VBD     play            play                            play
play            play
plays           VBZ     play            play                            play
plays           play
player          NN      player          player                          play
player          player
jumps           VBZ     jump            jump                            jump
jumps           jump
jumping         VBG     jump            jump                            jump
jump            jump
jumped          VBD     jump            jump                            jump
jump            jump
jumper          NN      jumper          jumper                          jump
jumper          jumper
running         VBG     run             run                             run
runn            run
ran             VBD     ran             ran                             ran
ran             run
runs            VBZ     run             run                             run
runs            run
runner          NN      runner          runner                          run
runner          runner
eating          VBG     eat             eat                             eat
eat             eat
ate             VBD     ate             ate                             at
ate             eat
eats            VBZ     eat             eat                             eat
eats            eat
eater           NN      eater           eater                           eat
eater           eater
talking         VBG     talk            talk                            talk
talk            talk
talked          VBD     talk            talk                            talk
talk            talk
talks           VBZ     talk            talk                            talk
talks           talk
talker          NN      talker          talker                          talk
talker          talker
singing         VBG     sing            sing                            sing
sing            sing
sang            VBD     sang            sang                            sang
sang            sing
sings           VBZ     sing            sing                            sing
sings           sing
singer          NN      singer          singer                          sing
singer          singer
writing         VBG     write           write                           writ
writ            write
wrote           VBD     wrote           wrote                           wrot
wrote           write
writes          VBZ     write           write                           writ
writes          write
writer          NN      writer          writer                          writ
writer          writer
swimming                VBG     swim            swim
swim            swimm           swim
swam            VBD     swam            swam                            swam
swam            swim
swims           VBZ     swim            swim                            swim
swims           swim
swimmer         NN      swimmer         swimmer                         swim
swimmer         swimmer
```

In [16]:

```python
import nltk
from nltk.stem import WordNetLemmatizer

# Initialize the WordNet lemmatizer
lemmatizer = WordNetLemmatizer()

# Sample surface forms of a word
surface_forms = ["running", "runs", "ran", "runner", "eaten", "ate", "eating",

# Perform morphological generation on the surface forms
print("Surface Form\tBase Form")
for surface_form in surface_forms:
    base_form = lemmatizer.lemmatize(surface_form, pos='v')
    print(f"{surface_form}\t\t{base_form}")
```

```
Surface Form    Base Form
running         run
runs            run
ran             run
runner          runner
eaten           eat
ate             eat
eating          eat
eater           eater
```

In [ ]:

In [ ]:

In [2]:
```
pip install statemachine
```

```
Collecting statemachine
  Downloading statemachine-0.1.tar.gz (1.4 kB)
  Preparing metadata (setup.py): started
  Preparing metadata (setup.py): finished with status 'done'
Building wheels for collected packages: statemachine
  Building wheel for statemachine (setup.py): started
  Building wheel for statemachine (setup.py): finished with status 'done'
  Created wheel for statemachine: filename=statemachine-0.1-py3-none-any.whl
size=1843 sha256=bcc43d62165dbeb1064c1e35b1faec57b4e8422cd585b899598808aae793
840b
  Stored in directory: c:\users\asus\appdata\local\pip\cache\wheels\51\98\63
\50f3917901b2239e5eb40f728ec73cb7403f50e81ca21a0691
Successfully built statemachine
Installing collected packages: statemachine
Successfully installed statemachine-0.1
Note: you may need to restart the kernel to use updated packages.

WARNING: Ignoring invalid distribution -treamlit (c:\users\asus\appdata\local
\programs\python\python310\lib\site-packages)
WARNING: Ignoring invalid distribution -treamlit (c:\users\asus\appdata\local
\programs\python\python310\lib\site-packages)
WARNING: Ignoring invalid distribution -treamlit (c:\users\asus\appdata\local
\programs\python\python310\lib\site-packages)
WARNING: Ignoring invalid distribution -treamlit (c:\users\asus\appdata\local
\programs\python\python310\lib\site-packages)
WARNING: Ignoring invalid distribution -treamlit (c:\users\asus\appdata\local
\programs\python\python310\lib\site-packages)
WARNING: Ignoring invalid distribution -treamlit (c:\users\asus\appdata\local
\programs\python\python310\lib\site-packages)
WARNING: Ignoring invalid distribution -treamlit (c:\users\asus\appdata\local
\programs\python\python310\lib\site-packages)

[notice] A new release of pip is available: 23.0.1 -> 23.1.1
[notice] To update, run: python.exe -m pip install --upgrade pip
```

In [ ]: