

CD Lab Practical 7

Name : Chetan Pardhi

Roll No. : 22

Batch : B1

Aim :

(A) Read the Research paper provided to understand the use of SDTS in language translation from English to Hindi.

(B) Write a program to implement tokenizer, syntax analyzer, and to perform syntax directed translation for conversion from English to Hindi as per the research paper. Also extend the procedure to generate correct Hindi sentence

Code:

```
import nltk

# nltk.download('stopwords')

from nltk.corpus import stopwords

from nltk.tokenize import word_tokenize, sent_tokenize

stop_words = set(stopwords.words('english'))

import sys

userInputText = "Harsh will eat an apple later"
```

```
engToHindiConverterDictionary = {
```

```
    "a": "एक",
```

```
    "an": "एक",
```

```
    "boy": "लड़का",
```

```
    "end": "है ",
```

```
    "ram": "ram",
```

```
    "apple" : "से ब",
```

```
    "eat" : "खाये गा",
```

```
    "later" : "बाद म"
```

```
}
```

```
proper_nouns = {
```

```
    "ram" : "राम",
```

```
    "harsh" : "हर्ष",
```

```
    "chandni" : "चांदनी"
```

```
}
```

```
token_sent = sent_tokenize(userInputText)
```

```
print(token_sent)
```

```
for sentence in token_sent:
```

```
    wordsList = nltk.word_tokenize(sentence)
```

```
    print(wordsList)
```

```
    wordsList = [w for w in wordsList if not w in stop_words]
```

```
    tagged = nltk.pos_tag(wordsList)
```

```
    print(tagged)
```

```

ansList = []

for wordNTag in tagged:

    word = wordNTag[0].lower()

    wordTag = wordNTag[1].lower()

    if word in proper_nouns:

        ansList.append(proper_nouns[word])

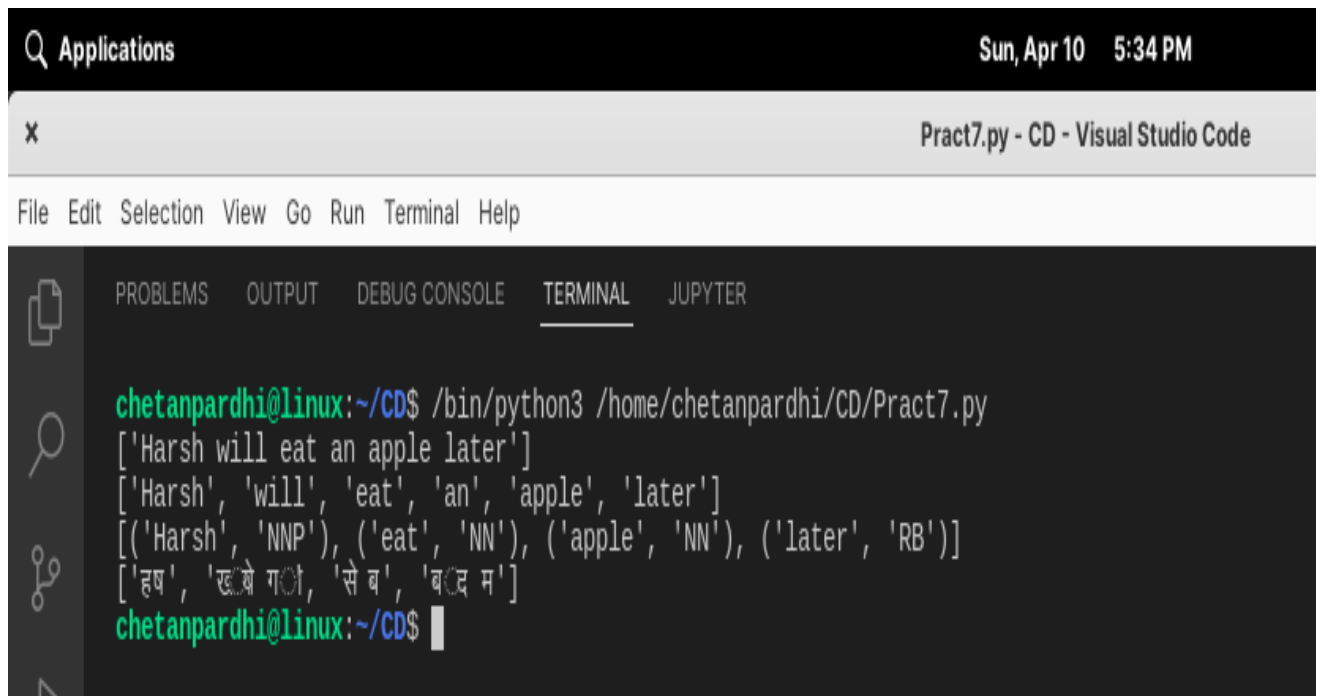
    else:

        ansList.append(engToHindiConverterDictionary[word])

print(ansList)

```

Screen-Shot:



The screenshot shows a Visual Studio Code window with a terminal running a Python script. The terminal output is as follows:

```

chetanpardhi@linux:~/CD$ /bin/python3 /home/chetanpardhi/CD/Pract7.py
['Harsh will eat an apple later']
['Harsh', 'will', 'eat', 'an', 'apple', 'later']
[('Harsh', 'NNP'), ('eat', 'NN'), ('apple', 'NN'), ('later', 'RB')]
['हर्ष', 'खे गी', 'से ब', 'बद म']
chetanpardhi@linux:~/CD$

```

