Devasy Patel

Practical 10

20BCE057

Date : 1/05/2023

AIM

- Implement the Soundex algorithm. Your code should generate four-character codes based on the pronunciation of English words.

Soundex rules

In [ ]:
```
rules = {
    **{ch: 0 for ch in 'AEIOUHWY'},
    **{ch: 1 for ch in 'BFPV'},
    **{ch: 2 for ch in 'CGJKQSXZ'},
    **{ch: 3 for ch in 'DT'},
    **{ch: 4 for ch in 'L'},
    **{ch: 5 for ch in 'MN'},
    **{ch: 6 for ch in 'R'}
}
```

In [ ]:
```
def step1(string):
    return [string[0].upper(), *(rules[ch] for ch in string[1:].upper() if ch in r
```

In [ ]:
```
step1('Hello World')
```

Out[ ]:
```
['H', 0, 4, 4, 0, 0, 0, 6, 4, 3]
```

In [ ]:
```
def step2(arr):
    ch = None
    ret = []

    for c in arr:
        if c == ch:
            continue
        else:
            ch = c
            ret.append(c)

    return ret
```

In [ ]:
```
step2(step1('Hello World'))
```

Out[ ]:
```
['H', 0, 4, 0, 6, 4, 3]
```

In [ ]:
```
def step3(arr):
    return [a for a in arr if a != 0]
```

In [ ]:
```
step3(step2(step1('Hello World')))
```

Out[ ]:
```
['H', 4, 6, 4, 3]
```

In [ ]:
```
def step4(arr):
    if len(arr) < 4:
        return arr + [0] * (4 - len(arr))
    else:
        return arr[:4]
```

In [ ]:
```
step4(step3(step2(step1('Hello World'))))
```

Out[ ]:
```
['H', 4, 6, 4]
```

SOUNDEX ALGORITHM

In [ ]:
```
def soundex(string):
    return ''.join([str(x) for x in step4(step3(step2(step1(string))))])
```

In [ ]:
```
soundex('Hello World'), soundex('helo wrld')
```

Out[ ]:
```
('H464', 'H464')
```