Sequences and Loops

Motivation: How to count the number of vowels in a word?

- You're given words like 'Boston', 'Wellesley', 'abracadabra', 'bureaucracies', etc.

- Tasks:
  - count the number of vowels in a word.
  - count the number of times a certain character appears in a word

```python
def countAllVowels(word):
    # body here

def countChar(char, word):
    # body here
```

Slides 8-3 to 8-12 explain what we need to know/learn to solve these problems.

Old friend: isVowel

```python
def isVowel(char):
    c = char.lower()
    return (c == 'a' or c == 'e' or c == 'i' or c == 'o' or c == 'u')
```

To think: How will the function isVowel be useful for solving our “counting vowels” problem?

Indices: accessing characters in a string

<table>
<thead>
<tr>
<th>In</th>
<th>Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td><code>word = 'Boston'</code></td>
</tr>
<tr>
<td>[2]</td>
<td><code>word[0]</code></td>
</tr>
<tr>
<td>[3]</td>
<td><code>B</code></td>
</tr>
<tr>
<td>[5]</td>
<td><code>o</code></td>
</tr>
<tr>
<td>[7]</td>
<td><code>s</code></td>
</tr>
<tr>
<td>[8]</td>
<td><code>word[3]</code></td>
</tr>
<tr>
<td>[9]</td>
<td><code>t</code></td>
</tr>
<tr>
<td>[13]</td>
<td><code>n</code></td>
</tr>
</tbody>
</table>

Notice
- 0, 1, 2, etc. are the indices (plural of index).
- Indices start at 0.
- Indices go from 0 to `len(word)-1`.
- We read `word[0]` as word sub 0.
- `[]` is known as the indexing operator.

To think: How will indices be useful for solving our “counting vowels” problem?
A possible solution: which is correct?

```python
word = 'Boston'
counter = 0
if isVowel(word[0]):
    counter += 1
if isVowel(word[1]):
    counter += 1
if isVowel(word[2]):
    counter += 1
if isVowel(word[3]):
    counter += 1
if isVowel(word[4]):
    counter += 1
if isVowel(word[5]):
    counter += 1
print counter
```

Does our solution work for all words?

- Do you think the right solution from 8-5 will work for all words: 'Wellesley', 'Needham', 'Lynn', etc.?
- What happens if we use an index that's greater than or equal to the length of the word?

```python
In [1]: word = 'Lynn'
In [2]: word[4]
IndexError: string index out of range
```

How to generate the correct indices of the string?

Properties of the `range` function

An optional third argument to `range` controls the step size between elements (which defaults to 1).

```python
In [1]: range(1, 10, 2)
Out[1]: [1, 3, 5, 7, 9]
```

To notice:
- With the help of the third argument of range, we can create different lists of integers.
- Step can be positive or negative.
- When step is negative, start value has to be larger than end value.
Introducing a new value type: lists

In [1]: type(range(0, 10))
Out[1]: list

We'll return to lists again in a few lectures. Lists are an example of "mutable" objects in Python, different from the "immutable" strings.

list() converts a string into a list of characters

In [2]: list("Wendy Wellesley")
Out[2]: ['W', 'e', 'n', 'd', 'y', ' ', 'W', 'e', 'l', 'l', 'e', 's', 'l', 'e', 'y']

We can also specify a list directly as a comma separated list of values

In [3]: phrase = ['A', 'lovely', 'a', 'day']
In [4]: phrase
Out[4]: ['A', 'lovely', 'a', 'day']

Sequences/Loops

Back to our vowel counting problem

word = 'Boston'
counter = 0
if isVowel(word[0]):
    counter += 1
if isVowel(word[1]):
    counter += 1
if isVowel(word[2]):
    counter += 1
if isVowel(word[3]):
    counter += 1
if isVowel(word[4]):
    counter += 1
if isVowel(word[5]):
    counter += 1
print counter

range solves our indexing problem, by generating the correct list of indices.

Try it out in the notebook to check that we get the same result.

Iterating Over Sequences with for Loops

One of the most common ways to manipulate a sequence is to perform some action for each element in the sequence. This is called looping or iterating over the elements of a sequence, and in Python is accomplished with a for loop.

for var in sequence:
    # Body of the loop
    statements using var

Generic form of a for loop

Concepts in this slide:
The combination of range and len to generate indices for a sequence.

Concepts in this slide:
New execution kind: iteration done through loops.

Concepts in this slide:
An example of a for loop with a conditional.
**for loop model example**

```python
word = 'Boston'
counter = 0
for i in range(len(word)):
    if isVowel(word[i]):
        counter += 1
print counter
```

To notice:
- There are three variables in the code: `i`, `word`, `counter`.
- The variables `i` and `counter` change values from one iteration to the next.
- Because `counter` is within a conditional, it's only updated when the condition is true.

---

**for loops without range**

- The `range` function provides a list of indices

**Mode 1**

```python
phrase = ['an', 'autumn', 'day']
for i in range(len(phrase)):
    print phrase[i] + '!
```

- We can also loop directly over any list. The code below produces the same output.

**Mode 2**

```python
phrase = ['an', 'autumn', 'day']
for word in phrase:
    print word + '!
```

---

**Strings and lists are both sequences**

```python
In [1]: word = 'Boston'
In [2]: word[2]
Out[2]: 's'
In [3]: len(word)
Out[3]: 6
In [4]: word + ' Globe'
Out[4]: 'Boston Globe'
In [5]: 'o' in word
Out[5]: True
In [6]: 'b' in word
Out[6]: False

In [1]: digits = range(7,13)
In [2]: digits[2]
Out[2]: 9
In [3]: len(digits)
Out[3]: 6
In [4]: digits + range(5)
Out[4]: [7, 8, 9, 10, 11, 12, 0, 1, 2, 3, 4]
In [5]: 10 in digits
Out[5]: True
In [6]: 15 in digits
Out[6]: False
```

A sequence is an “abstract” type, which serves as template for “concrete” types such as string or list.
Looping over a string

```
word = 'Boston'
counter = 0
for i in range(len(word)):
    if isVowel(word[i]):
        counter += 1
print counter
```

Can we avoid `range` in this code as we did in 8-15? It turns out yes, in the same way.

```
word = 'Boston'
counter = 0
for char in word:
    if isVowel(char):
        counter += 1
print counter
```

Operations in Sequences

<table>
<thead>
<tr>
<th>Operation</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>x in seq</td>
<td>True if an item of seq is equal to x</td>
</tr>
<tr>
<td>x not in seq</td>
<td>False if an item of seq is not equal to x</td>
</tr>
<tr>
<td>seq1 + seq2</td>
<td>The concatenation of seq1 and seq2</td>
</tr>
<tr>
<td>n*seq</td>
<td>n copies of seq concatenated</td>
</tr>
<tr>
<td>seq[i]</td>
<td>i'th item of seq, where origin is 0</td>
</tr>
<tr>
<td>seq[i:j]</td>
<td>slice of seq from i to j</td>
</tr>
<tr>
<td>seq[i:j:k]</td>
<td>slice of seq from i to j with step k</td>
</tr>
<tr>
<td>len(seq)</td>
<td>length of seq</td>
</tr>
<tr>
<td>min(seq)</td>
<td>smallest item of seq</td>
</tr>
<tr>
<td>max(seq)</td>
<td>largest item of seq</td>
</tr>
</tbody>
</table>

A simpler version of `isVowel` using `in`

```
# Old version
def isVowel(char):
    c = char.lower()
    return (c == 'a' or c == 'e' or c == 'i' or
c          c == 'o' or c == 'u')

# Simpler version
def isVowel(char):
    return char.lower() in 'aeiou'
```
The Slicing operator [:]

In [1]: word = 'Boston'
In [2]: word[2]
Out[2]: 's'
In [3]: word[2:5]
Out[3]: 'sto'
In [4]: word[:3]
Out[4]: 'bos'
In [5]: word[3:10]
Out[5]: 'ton'
In [6]: word[3:]
Out[6]: 'ton'
In [7]: word[0:6:2]
Out[7]: 'bso'
In [8]: word[::-1]
Out[8]: 'notsob'

How do indices work?

In [1]: digits = range(7, 12)
In [2]: digits[2]
Out[2]: 9
In [3]: digits[2:5]
Out[3]: [9, 10, 11]
In [4]: digits[:3]
Out[4]: [7, 8, 9]
In [5]: digits[3:10]
Out[5]: [10, 11]
In [6]: digits[3:]
Out[6]: [10, 11]
In [7]: digits[0:5:2]
Out[7]: [7, 9, 11]
In [8]: digits[::-1]
Out[8]: [11, 10, 9, 8, 7]

Indices in Python are both positive and negative. Everything outside these values will cause an IndexError.

In [7]: word[::-1]
   Out[7]: 'notsob'

This means: start at 0 until the end of sequence with step -1. And it works because of the negative indices.

Summary

1. Strings and lists are examples of sequences, ordered items that are stored together. Because they are ordered, we can use indices to access each of them individually and sequentially.
2. The indexing operator [] takes index values from 0 to len(sequence)-1. However, negative indices are possible too in Python, see 8-21.
3. If we can access each element of a sequence (string or list) one by one, then we can perform particular operations with them.
4. To access each element we need a loop, an execution mechanism that repeats a set of statements until a stopping condition is fulfilled.
5. When we loop over a sequence, the stopping mechanism is the arrival at the last element and not having where to go further.
6. We use the built-in function range to generate indices for sequences.
7. Python makes it easy for us to iterate over a sequence’s elements even without the use of indices. In fact we can write for item in sequence: and that will access each item of the sequence.
8. In addition to accessing one element at a time with [], one can use [:] (slicing) to get a substring or a sublist.