CIS 122

List Manipulations
Markov Text Generation

1. Corpus File → String → Word List
2. Word List → Markov Dictionary → Processed Word List
3. Processed Word List → Sentences → Generated Text
Lists Reviewed

- Lists are sequences of values
  - \( L = [1, 2, 3] \)

- These values can be of any type
  - \( L = [\text{True}, 2, 'C', (4.1, 5.2)] \)

- How many elements are in this list?
  - (If you're not sure, try the `len` function)
Lists Reviewed

- What can we do with lists?
  - Anything we can do with strings

- Index them
  - L[2]

- Add them
  - [1,2,3] + [4,5,6]

- Multiply them
  - [1,2,3] * 3
Lists Reviewed

● What can we do with lists?
  ○ Some things strings can't do

● Change them
  ○ L[2] = 100
  ○ Doesn't just reassign variable
  ○ Actually changes the list!

● Use list-specific methods
  ○ Like what?
Lists Reviewed

- Lists have a number of really useful methods

- Some return information
  - L.index('b')  # Return index of first 'b' in list
  - L.count('b')  # Return number of 'b' s in list

- Some just modify your list
  - L.append(x)  # Add x to the end of L
  - L.sort()  # Sort elements of L
List Quiz

What does myList look like after each line?

>>> myList = [ 10, 20, 30 ]

>>> myList.append(5)

>>> myList[ 0 ] = 15

>>> myList.sort()
List Quiz

What does this code do?

```python
L = []
for x in range(10):
    L.append(x)
print L
```
Markov Text Generation

- Let's write a function `processText(text)`
  - Takes string as input
  - Breaks string into list of words
  - Processes list of words to split out periods

- Where do we start?
Markov Text Generation

- Splitting a string is easy
  - Use the split method

- Processing the word list is harder
  - Need to iterate through, checking for periods
  - Build a new list as we go

- Give it a try