get\_sentence\_lists(text):

''' text is a string containing and a sentence. Return a list containing lists of strings; each list corresponds to each sentence, and contains the individual words in each sentence. All elements are returned lower case.'''

get\_sentence\_lists\_from\_files(filenames):

''' filenames is a list of strings, each one the name of a text file. Return (in order) a list of every sentence contained in all the text files in filenames. Format is the same as list returned from " get\_sentence\_lists(text)".'''

build\_semantic\_descriptors(sentences):

''' sentences is a list containing lists of strings representing sentences. Return a dictionary d. Every key, w, is a word that appears in one of the sentences, and has a corresponding value d[w], which itself is a dictionary representing the semantic descriptor of w (see assignment page).'''

most\_similar\_word(word, choices, semantic\_descriptors):

''' word is a string, choices is a list of strings, and semantic\_descriptors is a dictionary. Return an element of choices with the largest semantic similarity (computed using the above function) to word. If the similarity cannot be computer, return -1. In the case of a tie, return the element of choices with the smallest index. '''

run\_similarity\_test(filename, semantic\_descriptors):

''' filename is a string in the form of a file in the format .txt. semantic\_descriptors is a dictionary as described above. Return a percentage of questions on which most\_similar\_word() guesses the answer correctly using the semantic descriptors stored in semantic\_descriptors.'''

Test cases:

1. a)

THE CAT IS FRIENDLY.

ThE c'at i.s friendly

How's it going? It.s good.

.So, I hear St. George station is ... flooding.

Don'.t do that.

. . . .

. ' . '

h.e.l,l!o

That, was. fantastic!

1. b)

1. c)

Oranges are tasty.

Oranges. Oranges. Oranges.

Orange. ORANGES! Orange juice. Orange cake!

There was once a steak. Steaks are good. Steaks are good.

There is a lot of craze. Crazed mania. A lot of mania.

Hi. High. Higher.