# Blaðsíða 485-487 - Code listing 9.9-9.12

*def* add\_word(*word*, *word\_set*):

""" Add the word to the set. No word smaller than length 3. """

if len(word) > 3:

word\_set.add(word)

import string

*def* process\_line(*line*, *word\_set*):

""" Process the line to get lowercase words to be added to the set. """

line = line.strip()

word\_list = line.split()

for word in word\_list:

# ignore the '--' that is in the file

if word != "--":

word = word.strip()

# get commas, periods and other punctuation out as well.

word = word.strip(string.punctuation)

word = word.lower()

add\_word(word, word\_set)

*def* pretty\_print(*ga\_set*, *doi\_set*):

#print some stats about the two sets

print("Count of unique words of length 4 or greater")

print("Gettysburg Addr: {}, Decl of Ind: {}\n".format(len(ga\_set), len(doi\_set)))

print("{:15s} {:15s}".format("Operation", "Count"))

print("-"\*35)

print("{:15s} {:15d}".format("Union", len(ga\_set.union(doi\_set))))

print("{:15s} {:15d}".format("Intersection", len(ga\_set.intersection(doi\_set))))

print("{:15s} {:15d}".format("Sym Diff", len(ga\_set.symmetric\_difference(doi\_set))))

print("{:15s} {:15d}".format("GA\_DoI", len(ga\_set.difference(doi\_set))))

print("{:15s} {:15d}".format("DoI-GA", len(doi\_set.difference(ga\_set))))

# list the intersection words, 5 to a line, alphabetical order

intersection\_set = ga\_set.intersection(doi\_set)

word\_list = *list*(intersection\_set)

word\_list.sort()

print("\n Common words to both")

print("-"\*20)

count = 0

for w in word\_list:

if count % 5 == 0:

print()

print("{:13s}".format(w), end=" ")

count += 1

*def* main():

""" Compare the Gettysburg Address and the Declaration of Independece. """

gettysburg\_address\_set = *set*()

declaration\_of\_independence\_set = *set*()

gettysburg\_file = open("gettysburg.txt")

declaration\_independence\_file = open("declOfInd.txt")

for line in gettysburg\_file:

process\_line(line, gettysburg\_address\_set)

for line in declaration\_independence\_file:

process\_line(line, declaration\_of\_independence\_set)

pretty\_print(gettysburg\_address\_set, declaration\_of\_independence\_set)

main()