FUZZYWUZZY TEST REPORT

A potential bug has been found in the function partial_ratio().

Following is the tstl's final version of the test.

```
FINAL VERSION OF TEST, WITH LOGGED REPLAY:
ascii0 = 81                                                  # STEP 0
Coverage.py warning: No data was collected.
str10 = chr(ascii0)                                          # STEP 1
Coverage.py warning: No data was collected.
ascii0 = 74                                                  # STEP 2
Coverage.py warning: No data was collected.
str20 = chr(ascii0)                                          # STEP 3
Coverage.py warning: No data was collected.
str10 += chr(ascii0)                                         # STEP 4
Coverage.py warning: No data was collected.
str20 += chr(ascii0)                                         # STEP 5
Coverage.py warning: No data was collected.
assert (fuzz.partial_ratio(str10,str20) ==
fuzz.partial_ratio(str20,str10))  # STEP 6
ERROR: (<type 'exceptions.AssertionError'>, AssertionError(),
<traceback object at 0x7f1f210e3710>)
TRACEBACK:
  File "/home/eecs/Documents/CS562/Project/sut.py", line 1765, in
safely
    act[2]()
  File "/home/eecs/Documents/CS562/Project/sut.py", line 975, in
act41
    assert (fuzz.partial_ratio(self.p_str1[0],self.p_str2[0]) ==
fuzz.partial_ratio(self.p_str2[0],self.p_str1[0]))
```

The test basically is to check if partial_ratio() returns the same value even when arguments are swapped.

In the above test partial_ratio("QJ","JJ") is different than partial_ratio("JJ","QJ"). However, the ratio("QJ","JJ") = ratio("JJ","QJ"). This behavior occurs only when difflib is used. After installing the Levenshtein library, the same assertion hasn't failed yet.
My understanding is that the similarity of two strings should be the same irrespective of the order in which they are passed to the function. Also, edit distance commutes over 2 strings i.e., d(a,b) = d(b,a). Hence I consider this as a bug. However, the documentation doesn't mention it and so I need to thoroughly check what is going on in the function to validate the bug

Progress Report

Following are a few scenarios tested for all the four functions viz., ratio(), partial_ratio(), token_set_ratio(), token_sort_ratio().

1. Similarity between two empty strings is 0

2. Similarity between two equal strings is 100
3. Swap the arguments and check if the similarity is still the same (1 bug found)
4. fuzz.ratio() should behave in the same way as ratio() in difflib/levenshtein library
5. Partial ratio of two strings containing a common word is 100
6. Token sort ratio of 2 strings with same set of words arranged differently is 100
7. Token set ratio of 2 strings is 100 if second string contains at least one of the words from the first string

To be tested - extract() and extractOne()

Quality of SUT

The library is well written and has been tested.

Usability – The library functions are simple to use and are unambiguous. The code is readable with sensible comments.

Correctness – The functionality seems to be correct except when the arguments are swapped.

Testability – Testing the exact similarity between 2 strings is tricky and hence only absolute values like 0 or 100 is tested.

Coverage: Following is the coverage data from tstl.

```
STOPPING TEST DUE TO TIMEOUT, TERMINATED AT LENGTH 91
STOPPING TESTING DUE TO TIMEOUT
1.9293078056 PERCENT COVERED
3600.55668592 TOTAL RUNTIME
52495 EXECUTED
5249491 TOTAL TEST OPERATIONS
3480.81700945 TIME SPENT EXECUTING TEST OPERATIONS
81.3663294315 TIME SPENT EVALUATING GUARDS AND CHOOSING ACTIONS
4.13225769997 TIME SPENT CHECKING PROPERTIES
3484.94926715 TOTAL TIME SPENT RUNNING SUT
10.5344836712 TIME SPENT RESTARTING
0.0 TIME SPENT REDUCING TEST CASES
137 BRANCHES COVERED
105 STATEMENTS COVERED
```

Coverage report-

Name	Stmts	Miss	Branch	BrPart	Cover
sut.py	5259	4848	1220	0	6%
/usr/local/lib/python2.7/dist-packages/fuzzywuzzy-0.8.0-py2.7.egg/fuzzywuzzy/StringMatcher.py	52	36	18	0	23%
/usr/local/lib/python2.7/dist-packages/fuzzywuzzy-0.8.0-py2.7.egg/fuzzywuzzy/__init__.py	1	0	0	0	100%
/usr/local/lib/python2.7/dist-packages/fuzzywuzzy-0.8.0-py2.7.egg/fuzzywuzzy/fuzz.py	116	93	28	0	16%
/usr/local/lib/python2.7/dist-packages/fuzzywuzzy-0.8.0-py2.7.egg/fuzzywuzzy/string_processing.py	16	4	2	1	72%
/usr/local/lib/python2.7/dist-packages/fuzzywuzzy-0.8.0-py2.7.egg/fuzzywuzzy/utils.py	55	36	24	1	25%
Total	5499	5017	1296	2	7%