Chapter 5

Injecting faults into FuzzyWuzzy

5 faults were injected into FuzzyWuzzy to determine if the framework could detect each of the faults. This was done in two ways. The first was to inject individual faults into the project. The second multiple faults were injected. The same faults were injected in the same places in both cases. The only difference being in the first only one fault was injected at a time, in the second all the faults were injected at once.

The five injections were placed mainly in the function ratio, and in the function partial_ratio since a large number of other functions depends on them. Injection **1** was placed at line 50, the function would return 1 instead of 0 at this point. **2** was placed at line 53, we multiplied by 101 instead of 100. **3** at line 65, returned 1 instead of 0. **4** at line 91, we compared r to .005 instead of .995. **5** at line 96, we multiplied by 50 instead of 100.

During single fault injections, the project managed to run well enough except for in the specific cases where the fault injection occurred. The testing framework managed to detect that a fault was occurred in each case. During the multiple fault injection, FuzzyWuzzy became far more unstable. With each fault injected the ability to predict the outcome became more difficult. The testing framework managed to detect all the faults, but it was far more difficult to trace back the fault since there were multiple faults.

Initially, it was only planned to have the boundary test cases where a null string was being compared to a full string to fail. However, even after the multiple injections were implemented some of the boundary test cases were still passing. Specifically for tokenSetRatio, some of the test cases that were expected to fail did not fail, like TestCase3.

So even while being able to detect every fault, it is not guaranteed that the fault will be found even for test cases that were expected to fail. It is also possible that as projects become larger, the effect of a fault on a system could be extremely difficult to detect and find if the fault affects a large part of the system.