|  |  |
| --- | --- |
| Code Similarity Comparer(CSC), version 1.0  What the Comparer Does | |
| **Main Path** | **Alternate Paths** |
| 1. User submits a target source code file and an original source code file to the system  2. The system deletes all blank lines in the two files  3. The system scans through two files line by line and count the number of identical lines based on identity of plain text until the end of shorter one. The system calculate the percentage of identical lines in all lines in target file  4. The system output the percentage to user | 2.1 The system formats the two files  2.1.1 removes all comments  2.1.2 removes indentation and blank lines  2.1.3 converts duplicate whitespaces into single space  3.1 The system splits functions of two files and compares those functions by using longest common subsequence algorithm. The System calculates the similarity with the combination of function of highest similarity.  3.2 The system identifies and extracts all variable names from two files, and then using longest common subsequence algorithm to compare the two resulted files. |

|  |
| --- |
| User  Check similarity of two file  code-similarity |