

Plagiarism Detector Interface

Preprocessing

input: source file, example ('add.py')

output: fully processed version of source file

stripped version of source file, example ('stripped_add.txt')

HashFingerprinting

input: stripped version of source file, example ('stripped_add.txt')

output: fingerprints version of source file, example ('fingerprints.txt')

Winnowing

input: fingerprints version of source file, example ('fingerprints.txt')

output: prints a list of the fingerprints with their line numbers in the format (line, fingerprint)

Query - Perform a query of the original source code to be checked for plagiarism against database of source code.

input: (line, fingerprint)

output: list of matching fingerprints and lines and a list of their corresponding suspected plagiarized documents list of fingerprints and lines example (fp_pos1, fp_pos2)

where fp_posN = (line, fingerprint) of source

Visual

input: takes the original source file to be checked for plagiarism, the suspected plagiarized document, the list of each document's lines and fingerprints. example (source_file1, source_file2, fp_pos1, fp_pos2)