Exercise Set 4

More Python

University of Oslo - IN3110/IN4110

Fall 2021

Note: These exercises are not mandatory. You don't have to put your work into your Github repository.

superdiff

The standard utility `diff` takes two files as input, and outputs a file containing all changes which have to be made to the first file to make it into the second file. In this problem, you will create your own implementation of the standard `diff` utility. For convenience, your implementation should take two filenames from the command line, and treat the first as the "original" version of a file, and the second as a modified version. It should then go through the files line by line, and if a line has not been modified, print it with a `0` in front, if it has been added, print it with a `+` in front, and if it has been deleted, print it with a `-` in front. If a line has been modified, treat it as being a deletion of the original line, and then an addition of the modified line. Note that "line by line" is a bit misleading, as it is not enough to simply compare the two files line by line. Instead, you should try to find what the files have in common, to try to find the minimal modifications necessary to get the "modified file". There is however some ambiguity inherent to the problem. For example, if the original has the lines "A", "B" and the modified version has the lines "B", "A", was "A" deleted and later inserted, or was "B" inserted and later deleted? There are a lot of strange edge cases, so don't be too worried about making an implementation which handles absolutely all of them. You should try to make sure that your program is "reasonable" (i.e. if the original is "A", "B", "C", "D", "E", "F" and the modified file is "B", "C", "D", "E", "F", "G", your program should ideally output that "A" has been deleted and "G" has been inserted at the end, and not say "everything in the first file was deleted, then everything in the second file was inserted".