

TextDiff

[Jim's Pages](#) => [Java Pages](#) => TextDiff

I needed a program to compare two text files and report the differences - commonly called "diff". After a brief Google search I found a freeware [Diff by Ian F. Darwin](#). I liked the algorithm but the code didn't lend itself to producing the output I needed. So I took a shot at making a more generalized solution and wound up with a complete rewrite on the same algorithm.

The code is public domain, absolutely no warranties. Use and enjoy at your own risk. Please feel free to give it a shot and let me know how it works or any other comments or suggestions you have. Thanks!

Here's the source code: [TextDiff.zip](#). I have had trouble opening this zip file from a browser (Mozilla Firefox) but it seems to work fine if I download and then open. Let me know your experience!

TextDiff compares two text files or arrays of strings and generates a report of edit commands that would transform Old to New.

The algorithm (but no code) was taken from Java code by Ian F. Darwin, ian@darwinsys.com, January, 1997. Darwin's code was a translation of a C program by D. C. Lindsay, C (1982-1987)

The algorithm is:

1. For each each unique line of text create a symbol. The symbol state is: OldOnly, NewOnly, UniqueMatch (both files exactly once), or Other.
2. For each line, create a LineInfo object. Set state = symbol state and establish bidirectional links between UniqueMatch lines in the two files.
3. For each UniqueMatch in old create a "match block". Stretch match blocks forward and backward to include matching lines with any state, including other match blocks.
4. Build a Report of edit commands that can be used to tranform Old into New. Matching blocks generate match or move commands. Non-matching blocks generate insert, append, delete or change commands.
5. Iterate the commands to generate a report.

The DefaultReportWriter prints a human-friendly report to a PrintStream such as System.out. One could implement custom report writers to create machine-readable reports such as concrete editor commands.

Usage:

```
Report report = new TextDiff().compare( oldFileName, newFileName );
report.print( );
```
