

## CMPE 230 Systems Programming

### Project 2 (due April 26th)

( This project can be implemented in groups of at most two students. )

In this project, you will implement a file utility program called **filelist** that will traverse directories and report path names of files that satisfy some search criteria. The program will be implemented in Python. It will be invoked as follows on the console:

**filelist [options] [directory list]**

The arguments in [ ] are optional. If they are not given, the default action will be carried out. If no directory list is given, the default will be the current directory . If no option arguments are given, pathnames of all files will be printed by the program.

The following options will be available. Note that if multiple options are present, the files satisfying conditions of all options will be returned (i.e. option conditions are ANDed).

<b>-before datetime</b>	Files last modified before a date ( <b>YYYYMMDD</b> ) or a datetime ( <b>YYMMDDTHHMMSS</b> )
<b>-after datetime</b>	Files last modified after a date (YYYYMMDD) or a datetime (YYMMDDTHHMMSS)
<b>-match &lt;pattern&gt;</b>	Filenames matching a Python regular expression <b>&lt;pattern&gt;</b>
<b>-bigger &lt;int&gt;</b>	Files having sizes greater than or equal to <b>&lt;int&gt;</b> bytes. Note that <b>&lt;int&gt;</b> can also be given as kilobytes, megabytes and gigabytes, for example as, 2K, 3M, 7G.
<b>-smaller &lt;int&gt;</b>	Files having sizes less than or equal to <b>&lt;int&gt;</b> bytes. Note that <b>&lt;int&gt;</b> can also be given as kilobytes, megabytes and gigabytes, for example as, 2K, 3M, 7G.
<b>-delete</b>	The files should be deleted.
<b>-zip &lt;zipfile&gt;</b>	The files should be packed as zipfile.
<b>-duplcont</b> <b>-duplname</b>	Note that only one of <b>-duplcont</b> or <b>-duplname</b> can be given. <b>-duplcont</b> : Files whose contents are the same should be listed. The listing should be in sorted order with a line of ----- characters separating the duplicate sets of files. <b>-duplname</b> : Files whose names are the same should be listed. The listing should be in sorted order with a line of ----- characters separating the duplicate sets of files.
<b>-stats</b>	Traversal statistics should be printed at the end of files listing output. The statistics will be as follows: total number of files visited, total size of files visited in bytes, total number of files listed, total size of files listed in bytes. If <b>-duplcont</b> option is given, additionally, the following information will be printed: total number of unique files listed, total size of unique files in bytes. If <b>-duplname</b> option is given, additionally, the following information will be printed: total number of files with unique names.
<b>-nofilelist</b>	No file listing will be printed

**Grading**

Your project will be graded according to the following criteria:

Documentation (written document describing how you implemented your project)	12%
Comments in your code	8%
Implementation and tests	80%

**Late Submission**

If the project is submitted late, the following penalties will be applied:

- $0 < \text{hours late} \leq 24$  : 25%
- $24 < \text{hours late} \leq 48$  : 50%
- $48 < \text{hours late} \leq 72$  : 75%
- $\text{hours late} > 72$  : 100%