Code-a-thon Project Idea - TextFinder

Jim Fawcett, 8/18/2020

Here’s a problem definition I thought about during our Zoom meeting. It approximates some of the suggestions and is quite different from others. Intended to be a strawman.

1. Goals:
   1. Clear and simple statement
   2. Rich enough to provide interest to students from different experience levels
   3. Easily testable program results
   4. Can implement in many common programming languages: C, C++, Python, Java, …
2. Proposed coding problem: Text Finder
   1. Identify all files in a directory subtree that contain a specified text.
   2. Specify root path, file pattern(s), and search text on command line
   3. Display file name and path without duplication of path name, e.g., organized by directory
   4. Interesting extensions:
      1. Select files by pattern, e.g., \*.h,\*.cpp
      2. Replace text by regular expressions for both search text and file names
3. Required student activities:
   1. Develop software design (could be a pre-thon event, say 24 hours, with mentor support)
      1. Identify relevant program tasks:
         1. Parse command line
         2. DFS traversal of directory tree
         3. Open matching files
         4. Search for text
         5. Display results in some meaningful order
      2. Develop software structure
         1. Parts (call graph or packages or classes)
         2. Responsibilities of each part
         3. Execution time-line
         4. Potential for forking execution with threads or processes (not implemented)
      3. Define error handling
         1. Parsing errors, openiing errors, reading errors
      4. Define tests
         1. Empty directory
         2. Directory tree with specified files, some matching pattern, some not matching
   2. Implement TextFinder
      1. Credit for partial results weighted by experience.
      2. Run on directory contents provided by Code-a-thon directors
         1. Contains some folders with binary files
         2. Only one correct result known by Directors
   3. Present design, relate that to displayed results