csce215 — UNIX/Linux Fundamentals Spring 2022 — Assignment 6

This assignment is intended to provide some practice and additional content for the material covered in lecture on Monday, March 14. You'll practice using find, grep, and regular expressions. The assignment is meant to be started in the lab sessions on Wednesday, March 16 and Thursday, March 17. It must be submitted by 11:59pm on Sunday, March 20. A total of 92 points are available.

1 Get started

First, a reminder: If you are in the lab during the scheduled time, don't forget to use the QR code or the password to claim credit for attendance.



As usual, you should use recbash to record your terminal session. But there is no need to create an assignment6 directory nor to copy any files into your home directory.



2 Exploring a big program: Finding files

First, navigate to /class/215/assignment6/xfig-3.2.8b. There you'll find the source code, mostly in C, for a recent version of a classic drawing program called xfig. This includes about 3500 files spread across over 100 directories. We'll use this as a representative example of the source code for a 'real' piece of software —rather than a small program designed specifically for a class— that has been developed over a long period of time and used by thousands of people around the world.

Let's use this directory to practice searching for certain files. For each of the following criteria, use find, with appropriate tests, to print a list of exactly the files that match those criteria. *Some of these will require tests that we did not directly discuss in the lecture, so you'll need to check the documentation for the* find *command to complete these.*

- 1. Files with an .xpm extension. Those are images in a certain format that traditionally worked well with graphical Linux programs. (Hint: There are 2 of them.)
- 2. Directories whose name contains the word 'metric'. (Hint: There are 3 of them.)
- 3. Files whose size is exactly 351 bytes. (Hint: There are 13 of them.)
- 4. Files with a .c extension whose file size is at least 90,000 bytes. (Hint: There are 4 of them.)
- 5. Files that are executable. (Hint 1: There are 8 of them. Hint 1: There's a specific test that asks find to check whether something is executable. Hint 2: Be sure to exclude directories, which may also be considered 'executable' We'll discuss permissions, including this detail, in Chapter 7.)
- 6. Files (not directories) that are newer, i.e. have a modification time more recent than, the format documentation file doc/html/fig-format.html. (Hint: There are 37 such files.)

From within the xfig-3.2.8b directory, give a find command that lists exactly the files and/or directories that meet each of the criteria listed above. 40 points

3 Exploring a big program: Searching for patterns

Now let's look for some patterns inside the files, to practice searching for various types of code. For each of the patterns below, use grep with its -r and -E options to find and print lines, from any file in xfig-3.2.8b, that match those patterns.

- 1. Lines that contain the string ellipse, in any context. (Hint: There are 995 such lines.)
- 2. Lines that mention one of xfig's 'drawing initialization' functions. These are functions whose names start with init_ and end with _drawing. (Hint: There are 58 such lines.)
- 3. Lines that mention one of the 'unsigned integer' data types uint16, uint32, uint16_t, or uint32_t. (Hint: There are 22 such lines.)
- 4. Lines that contain either 'fixme' or 'todo', either of which are sometimes added to comments as a reminder that the code is still a work-in-progress. These can be either upper case or lower case. (Hint: There are 33 such lines.)

- 5. Lines that *begin with* 4 2.¹ (Hint: There are 193 such lines.)
- 6. Lines that *end with* a pound sign (#) followed by exactly 6 hexadecimal digits (that is, 0-9, a-f).²

(Hint: There are 3464 such lines.)

From within the xfig-3.2.8b directory, give a grep command that lists exactly the source code lines that meet each of the criteria listed above. 40 points

4 Exploring a big program: What's left of the copyright

Finally, here's one more grep exercise that will require a little more care. Let's look for lines that contain a copyright notice. Each copyright notice consists of the word 'copyright' (which may or may not be capitalized), possibly followed by a '(c)' (which may or may not be capitalized and might be missing completely), followed by a four-digit year. To give you a sense of how varied these lines might be, here are some examples from this xfig code.

```
# Copyright (C) 2004-2020 Free Software Foundation, Inc.
> Parts Copyright (c) 1991 by Paul King
uses code from xpaint (Copyright 1993 David Koblas (koblas@netcom.com)
and Copyright 1995, 1996 Torsten Martinsen (bullestock@dk-online.dk))
<br>Parts Copyright (c) 1989-1999 by Brian V. Smith
* Parts Copyright (c) 1991 by Paul King
```

Just like the exercises above, use grep with a regular expression to find all of the lines with these kinds of copyright notices. As a hint, there are 884 such copyright notices.



¹In case you are quite reasonably asking: 'Who cares?' In the xfig file format, diagrams are stored as text files, in which (more or less) each line defines a object in the diagram. Lines that start with 4 represent text labels; the 2 after that indicates that the label should be right-justified. So you in this case, you are searching for right-justified text labels.

²Who cares? In the xfig file format, these lines represent are used to represent custom colors to use for other objects in the diagram.

5 It's Done

When you've completed all of the tasks above, submit as usual. End the recbash recording using Ctrl-D or exit. Check its transcript to ensure that it shows all of the steps mentioned in the green boxes. Then upload and submit.



