ASSIGNMENT 5

The purpose of this assignment is to introduce you to data carving.

Description

In this assignment, you will write a C/C++ program called *ImageCarver* that extracts (carves) JPEG images from a file representing the unallocated space found on a disk.

You are free to choose and implement one, or a mix, of the data carving methods from the following list.

1. Header-footer carving
2. Header-maximum length carving
3. File structure carving

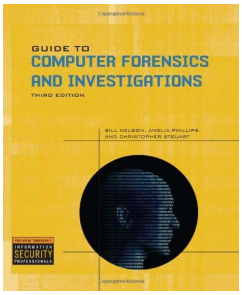
Output

For every image you carve, the following output must be produced:

- A `.jpg` file of the image. You need to copy the carved bytes to a file. Name the files using numbers – first file is `1.jpg`, then `2.jpg`, and so on.

- Print to the display (one line for each image):
  - name: the file name
  - start: offset in *unalloc.img* where you found the beginning of the image (the first byte)
  - size: number of bytes carved

For example, *unalloc.img* has the following image starting at offset `0x281000` and ending at offset `0x288929` (both inclusive). This image has a size of `31018` bytes.



Your program should not output the same image as two different files unless they have been found in two different locations. Hence, if you choose to implement multiple methods, make sure that the output is a combined output without duplicates.

Submission

Follow file naming conventions, and upload the *ImageCarver.cpp* (or *.c*) file to Canvas. You must fully comment your code (you may provide a README file in addition) so that the method you are implementing becomes clear from the comments.

Grading

The assignment is worth **100 points**. 60 of these points will be given based on the number of pictures your program is able to carve out, and how close the size of carved images are to that of the real ones. Remaining 40 points: 20 points for correctly displaying output and 20 points for clean and commented code.

How will the 60 points be assigned?

Points will first be assigned as follows.

- You will receive 5 points for each unique image that you carve. The image must be viewable in an image viewer, either fully or partially, to be eligible for these points. A thumbnail inside a JPEG image should be treated as a separate image. If your program reports multiple images with the same start offset, the first of these images reported in the printed output will be used to determine eligibility.
- 2 points will be deducted for each duplicate image. A duplicate is not an image that looks the same as another (thumbnails will look the same as the main image), but one whose start offset is same as another reported image.
- 2 points will be deducted for every unique image whose reported size is 50KB or higher more than its actual size.

Your total points will then be divided by the maximum possible points (not revealed now), and then multiplied by 60. This will be your total points out of 60. A negative score will be converted to zero.

A program that does not compile is a program that you did not submit at all. Remember the GTA is not required to debug your program to give you partial points.

The late policy is available in the course syllabus. **You must work alone on this assignment.**