Design Document - Getting Acquainted with Unix and C

Lev Teytelman

January 8, 2022

1 Program Description

This script compiles, runs, and uses the results of collatz.c to create graphs displaying various stats about the Collatz sequence. It runs the program about 10000 times with inputs from 2 to 10000, inclusive, and stores the results in an array. It then uses the results to graph the sequence lengths, max values, and a histogram of the values using gnuplot.

2 Included Files

- plot.sh - The bash script used to compile and run collatz.c and plot the results.
- collatz.c - The provided C file containing the code to print the Collatz sequence.
- Makefile - The file used to create more readable aliases for removing old binaries and compiling new ones.
- README.md - Contains instructions for running the program.
- DESIGN.pdf - This document, describes the project, its files, and objectives.
- WRITEUP.pdf - Contains a summary of the results, including the plots and an analysis of the script.

3 Structure

The script starts by creating the `/tmp/` directory, followed by removing old files and building the `collatz.c` file. After this, two arrays are created named `lengths` and `maximums`. The script loops through all numbers from 2 to 10,000 (inclusive) and runs the newly-compiled C file with the incrementing number as the input.

The output is then put into an array called `nums`, which is then used to populate the other two by taking the length of the array and appending it to `lengths` and sorting the array and appending the largest value to `maximums`.

After the for loop completes, the arrays are saved to temporary files `/tmp/lengths.dat`, `/tmp/maximums.dat`, and `/tmp/length_hist.dat`.

After creating the files, the script feeds `gnuplot` a series of inputs, first setting various properties about the environment and eventually plotting the `.dat` files onto three different PDFs.