

Assignment One Design.pdf

By James Yim

Description of Program:

This program is a script which will allow us to plot certain things pertaining to a Collatz sequence. The script will take a Collatz sequence and produce plots based on the length of the sequence, the maximum collatz sequence value, and a histogram of the Collatz values. The script will look at every Collatz sequence from 2 to 10,000 and build a graph.

Files to be included in asgn1

- Plot.sh
 - This script creates plots based on Collatz sequences given by .collatz.c. This file will plot the lengths of sequences, plot maximum sequence values, and make a histogram based on sequence lengths.
- Collatz.c
 - This file was given and contains the implementation of the Collatz sequence values.
- Makefile
 - This file was given and directs the compilation process of the Collatz sequence program which is collatz.c
- Readme.md

- This file is in markdown format. It describes how to use the script which is plot.sh and describes Makefile.
- Design.pdf
 - This document describes the design and the design process. It will contain pseudocode and explanations on how the program works.
- Writeup.pdf
 - This will include the plots that the script has made. It will also talk about the bash scripting commands that were used.

Design & Pseudocode

-For plotting sequence lengths:

Loop through 1 to 10,000 for the amount of iterations

Take a collatz sequence and count how many lines and put it in a file for gnuplot to read

Take the iteration number and put it in the file for gnuplot to read for coordinates

Plot the sequence using gnuplot

-for plotting max value

Loop through 1 to 10,000 for the amount of iterations

Take a collatz sequence and record the largest value and put it into a file for gnuplot

Plot the sequence using gnuplot

-for plotting the histogram

Loop through 1 to 10,000 for amount of iterations

Get lengths of collatz sequences

Take the lengths of sequences and organize them in increasing order with repeating numbers next to each other

Take the file with the sorted numbers and count how many times each number repeats

Plot the sequence using gnuplot with the x value of the plot as the collatz sequence and the y value how many time it repeats.

Notes about the pseudocode/structure

- `Rm -f <filename>`
 - This line is used in the beginning of each plotting sequence.
 - It is used in order to clear the file for the script to be ran multiple different times
- `Echo -n $i`
 - This is used to get coordinates for gnuplot to plot.
 - Used as the x axis in lengths plot and max sequence plots.
- `./collatz -n $i | wc -l >> <filename>`
 - This takes collatz numbers and counts how many numbers there are by linecounts and puts it into a file
- `./collatz -n $i | sort -n | tail -n 1 >> <filename>`
 - This takes all the collatz numbers in a specific sequence and sorts it from smallest to largest. The tail part gets the last number which is the largest number and puts it into a file
- `Cat /tmp/lengthsInOrder.dat | uniq -c >> <filename>`
 - This takes the lengths in order, and takes the sorted numbers and counts how many times each number is present.

Errors encountered

- There was an error where I was not able to get a space in between the two coordinates for gnuplot to read
 - I fixed this by adding a backslash.
- There was another issue where when trying to get unique values it would not count values that were not next to each other.
 - In order to fix this I sorted the lengths first and then sorted by uniq

Credit

- I attended Eugene's section on 1/7/22. He helped with how to do lengths plotting. Also gave general advice on how to go about the assignment.