Assignment 2

In this project we are assigned to find and print duplicate files or directories. I used some built in functions such as os.walk, hashlib.sha256, os.listdir, os.path.join, os.path.basename etc.

In the first part of my code I am parsing the given command line arguments with arge.parse. Setting default values of flags are handled in this part. Second part is determining whether this command wants to print duplicate files or directories. According to -d and -f flags this process starts. To find duplicate directories I used recursion. I wrote three similar functions. These functions are for -c -n and -cn flags. There are minor changes in these functions. So, they may seem same at first look.

traverse\_dir\_for\_C(direktory, setofDirs,sizeofDirs)

this function for -c flag. It hashes the contents of files and directories recursively.

traverse\_dir\_for\_N(direktory, setofDirs)

this function for -n flag. It hashes the names of files and directories recursively.

traverse\_dir\_for\_CN(direktory, setofDirs)

this function for -cn flag. It hashes the names and contents of files and directories recursively.

For finding duplicate files I used os.walk function not recursion. By doing this I get all files easily and hash them immediately. Finding duplicate files are easy but printing them under some restrictions is hard. If -s flag is set, then I need to print duplicate files based on their sizes. If not, it must be printed in alphabetical order.