## **Домашнее задание 2**

Поиск дубликатов файлов.

Скрипт получает в качестве параметра командной строки имя каталога. Требуется обойти рекурсивно этот каталог вы найти все файлы с идентичным содержанием. Символьные ссылки а также файлы, имена которых начинаются с . или заканчиваются на ~ должны игнорироваться.

Каждая строка вывода программы должна содержать имена файлов, разделенные символом ‘:’, с идентичным содержимым. Если содержимое какого-то файла никем не дублируется выводить его не нужно.

Например, у нас есть следующая иерархия:

top\_dir/

+- file1

|

+- file2

|

+- .hidden

|

+-dir1/

| +- file1

| |

| +- file3

|

+-dir2/

+- file1

|

+- file3

|

+- .hidden

Мы передает top\_dir в качестве параметра командной строки скрипту. Кроме того, считаем что файлы с одинаковым именем (file1, file2 или file3) имеют одинаковое содержимое, тогда вывод программы может быть следующим:

file1:dir1/file1:dir2/file1

dir1/file3:dir2/file3

Порядок строк и порядок файлов внутри строк не важен. Имена файлов могу выводиться либо как в примере относительно top\_dir (не включая ее), либо в виде абсолютного пути. Обратите внимание, что file2 есть только в одном экземпляре, поэтому в результате программы он отсутствует. Имя файла .hidden начинается с точки, поэтому в выводе программы он также отсутствует.

При выполнении задания окажутся полезными модули sys, os и hashlib.

### Сроки

Мягкий дедлайн для этого задания 28.09 23:55 по МСК. Задания **принятые** до этого строка оцениваются максимум в 3 балла. Жёсткий дедлайн для этого задания 05.10 23:55 по МСК. Задания принятые после мягкого дедлайна и до этого срока оцениваются максимум в 1,5 балла. Все не принятые или принятые после этого срока задания оцениваются в 0 баллов.

UPD: Задания, не засчитанные до мягкого дедлайна необходимо присылать с тестами.