Kabeer Ahmed SE-19028

LAB 5

Exercise:

1. Are the data (128 bytes) generated by md5collgen completely different for the two output files? Please identify all the bytes that are different.

In general, it is highly unlikely that the data generated by the MD5 hash function for two different files will be completely identical. MD5 is a cryptographic hash function, which means that it is designed to produce unique hash values for a given input. Even a small change in the input will result in a vastly different output.

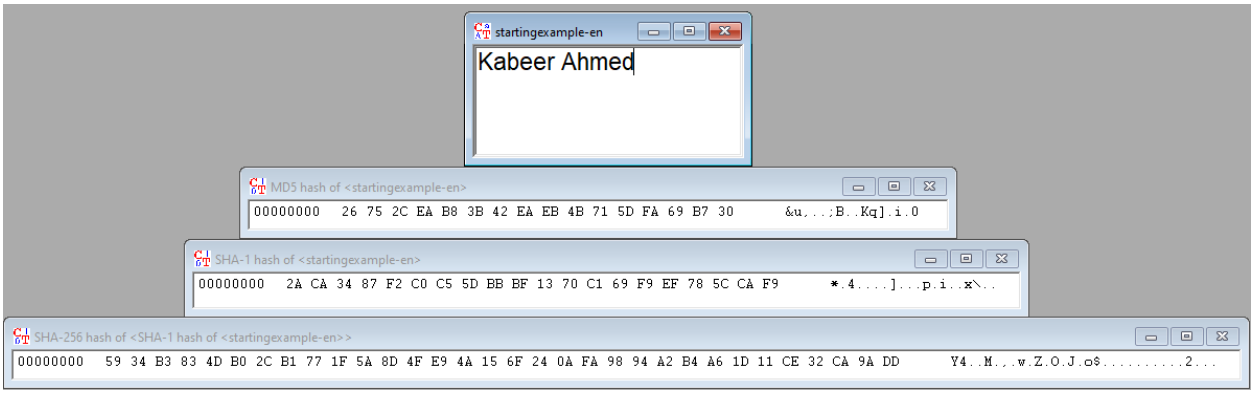
However, it is possible that two different files will produce the same MD5 hash value, a phenomenon known as a "hash collision." This is because the MD5 hash function has a limited output space of 2^128 possible values, and it is possible that two distinct inputs may produce the same output.

If you want to know the specific bytes that are different in the two output files, you would need to compare the two files byte by byte, and record the positions at which the bytes differ.

2. If the length of your prefix file is not multiple of 64, what is going to happen?

If the length of the prefix file is not a multiple of 64 in the MD5 algorithm, the remaining bytes will be padded with zeros until the length of the prefix file is a multiple of 64. This ensures that the data can be processed properly by the algorithm's block size.

3. Calculate Hash or fingerprint of your name using MD5, SHA-1 and SHA-256 algorithms and explain the result?



Kabeer Ahmed SE-19028

4. The hash compute for any document the output is a constant length, independent whether if the input length is 1 byte or 1 Mbyte. Explain how the hash algorithms achieve this fixed length?

Hash algorithms, such as MD5, use a process called "hashing" to convert an input (or "message") of any length into a fixed-length output, known as a "hash value" or "digest." This process typically involves several steps, such as:

- Breaking the input into blocks of fixed size.
- Applying a mathematical function, such as a compression function, to each block.
- Combining the resulting hash values from all the blocks to produce a final hash value.

The compression function used in MD5 is a combination of logical operations such as bitwise operations, modular addition, and logical shifts. It uses a fixed set of mathematical operations and a fixed set of constants, which are defined in the algorithm.

In summary, the fixed length of the output is achieved by applying a fixed mathematical function to each block of the input, and then concatenating the resulting hash values.