

Assignment 7

Exercise 1

Hash functions must provide consistent output for the same input, as failing to do so would result in losing the ability to locate the item after it has been placed in the hash table.

Out of these hash functions, which ones exhibit consistency?

- a. $f(x) = 1$ (Returns "1" for all input) 0.5 point
- b. $f(x) = \text{rand}()$ (Returns a random number every time) 0.5 point
- c. $f(x) = \text{next_empty_slot}()$ (Returns the index of the next empty slot in the hash table) 0.5 point
- d. $f(x) = \text{len}(x)$ (Uses the length of the string as the index) 0.5 point

a. $f(x) = 1$ (Returns "1" for all input).

Is consistent since it always returns "1", although it will create collisions if there is more than one input.

b. $f(x) = \text{rand}()$ (Returns a random number every time). \ Is not consistent since it returns a random number every time, making it impossible to map back.

c. $f(x) = \text{next_empty_slot}()$ (Returns the index of the next empty slot in the hash table)

Is not consistent since it returns the current state value of the hash table, not the input value.

d. $f(x) = \text{len}(x)$ (Uses the length of the string as the index).

Is consistent because it returns the length of the input.

So only hash function a. and d. are consistent.

Exercise 2

Exercise 2 (3 points) : Having a good distribution is a crucial factor for hash functions. Their goal should be to map items as uniformly as possible across the hash table. The most unfavorable scenario occurs when a hash function maps all items to the same slot in the hash table.

Consider these four hash functions that operate on strings:

A. Returns "1" for all input.

B. Uses the length of the string as the index.

C. Uses the string's first character as the index, resulting in all strings starting with the same letter being hashed together.

D. Maps each letter to a prime number, such as $a = 2$, $b = 3$, $c = 5$, $d = 7$, $e = 11$, etc.

The hash function for a string is the sum of all the characters, with the result modulo the hash's size. For instance, if the hash size is 10, and the string is "bag," the index is calculated as:

$$(3+2+17)\% 10=22 \% 10 = 2.$$

Which hash functions would provide a good distribution for each of these examples?
Assume a hash table size of 10 slot.

2.1 A phonebook where the keys are names and values are phone numbers. The names are as follows: Esther, Ben, Bob, and Dan. (1 point)

D. Since it uses prime numbers and the modulo operator to map each letter to a number.

2.2 A mapping from battery size to power. The sizes are A, AA, AAA, and AAAA. (1 point)

B. Since the characters are the same and the only difference is its string length.

2.3 A mapping from book titles to authors. The titles are Maus, Fun Home, and Watchmen. (1 point)

D. Same as in the phonebook question. This hash function gives the best distribution from the given hash functions.

Exercise 3 (1 point) :

A hash function uses the following formula to compute the hash value for a given input string:

$$\text{hash_value} = (\text{sum of ASCII values of each character in the string}) \% 13$$

What is the hash value for the input string "hello"?

ASCII value for each letter in "hello":

$$h = 104$$

$$e = 101$$

$$l = 108$$

$$l = 108$$

$$o = 111$$

Hash sum is 532

$$\text{hash_value} = 532 \% 13 = 12$$