Course: Data Structures (CSE CS203A, 114-1)

Study Note: Hashing

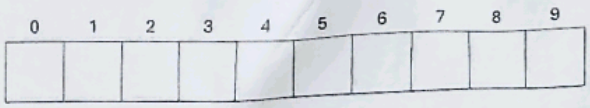
Student ID: 113346 Student Name: 趙暐勛

Definition

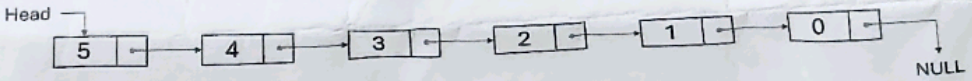
1. Hash Function - A hash function is a function that converts a key (such as a number or string) into an integer index. This index is then used to determine where the data will be stored in a hash table.
2. Hash Table - A hash table is a data structure that stores key–value pairs. It uses a hash function to compute an index in a data structure (called a bucket) where the data will be placed.
3. Collision Handling - A collision occurs when two or more keys are assigned to the same index by the hash function. Collision handling refers to the methods used to store and retrieve these multiple items that share the same index.

Data Structures: Visualization

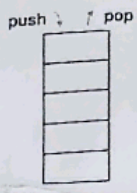
(1) Array



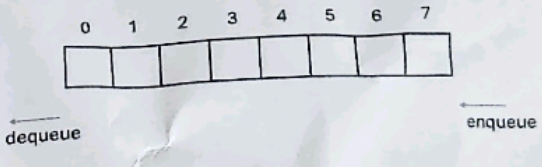
(2) Linked List



(3) Stack

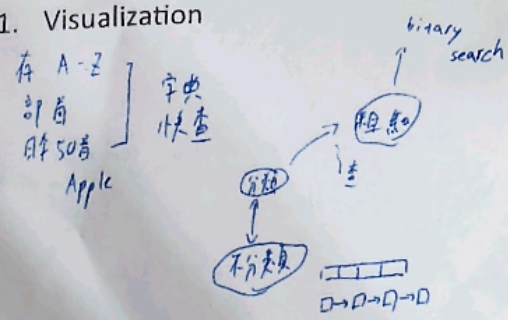


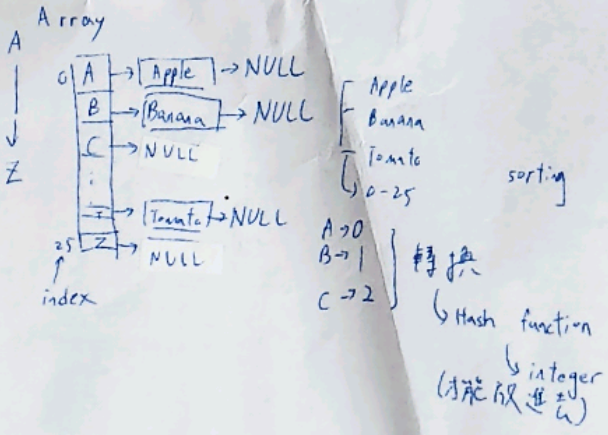
(4) Queue



Note

1. Visualization





2. Abstract Data Type

Definition:
No two pairs have the same key
several pairs with the same key

sorting → order
Effective

A→0
B→1
Z→25
Hashing function

data deposite

Hash Table ⇒ Array + Linked List

index ← Hash function

0 ← Hash (Apple) ⇒ Ascii(A)-65=0
1 ← Hash(Banana)

Collision
key
value ⇒ key-value pair

for example:
Array {23, 33, 43, 53, 63, 73}
Hash = input mod 3
23, 33, 43, 53, 63, 73
24, 34, 54

① space ⇒ realloc
② capacity → realloc
Hash function → key diverse → collision / collision

3. Implementation

Dictionary : characteristic
(collection of word)

2D Array :

Underlying structure
- Array : used for bucket storage
- Linked list : handles collisions

Hybrid : Array + linked list
Array index = bucket (from hash function)

Real-world Application : DNS Caching

collision = different keys, same hash address
probing - collision-resolution technique
open addressing (entire array implementation) - linear probing / quadratic probing / double hashing
open address

chaining

key mod n
bucket (index)

Large Array

key → index
key mod M → num