Study Note: Hashing

Student ID: 113428 Student Name: 徐妍甄

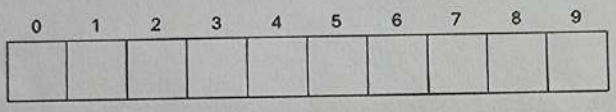
Definition

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.

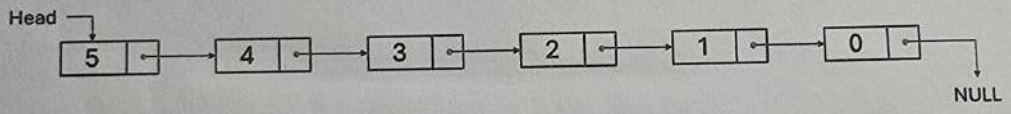
分類→不分類
ex: A-Z ····· Apple
香蕉 ····· 卜
⋮

Data Structures: Visualization

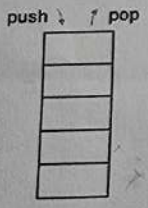
(1) Array



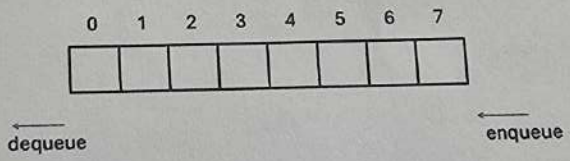
(2) Linked List



(3) Stack



(4) Queue



Note

1. Visualization

array + linked list

A → Apple → null
B → Banana → null
⋮
Z → Zoo → null

hash Table

A→0
B→1
⋮
Z→25
⇒ hash function

ex:
f(Apple) = 0
f(banana) = 1
⋮
f(zoo) = 25

Collision (碰撞)

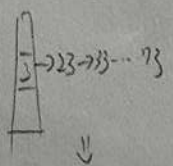
hash function(x) = 0
(y) = 0
(z) = 0

ex: 13, 33, 43, 53, 63 ?3

function : %10

3 3 3 3 3 3 ⇒ collision!

①



⇓

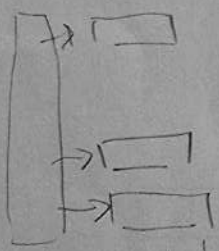
② 重設計 hash function (key 不夠 diverse)

2. Abstract Data Type

3. Implementation

① need Index ⇒ array

② array 上的 ⇒ 空間利用率太低



array + linked list

③ key 利用 hash function 打散 (避免 collision)

· Hash Table → use prime size, avoid repeating patterns (?)

④ All array

after hash 的 data ↑probing：找下一個可以存 data 的空間

a → 3

b → 3 → 4

c → 3 → 5

→ primary collision :(

不斷找資料

- linear probing →
- quadratic probing
- Double hashing

空間利用最大化 (吧)