

Student ID: 1131523

Student Name: 蘇茂瑄

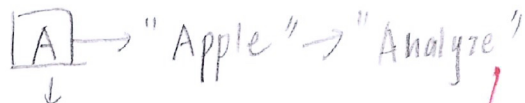
100

Part A: Hash Table Definitions (Conceptual Understanding)

Q1. Define "collision" in the context of hash tables.

A1: 碰撞, 就是有多個 key 被分配給同一個 index

Q2. What is a "bucket" in a hash table?

A2: 
bucket. (儲存位置/儲存名稱)

Q3. Define "load factor (α)" and explain why it affects performance.

A3: load factor 負載因子.

$$\Rightarrow \frac{\text{elements}}{\text{total buckets}} \Rightarrow \alpha \text{ 越高} \Rightarrow \text{collision 越多.}$$

Q4. What is "primary clustering," and which probing method suffers from it?

A4: primary clustering 初次聚集 linear probing.

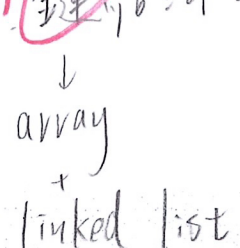
\Rightarrow open addressing 使有一大串連續的佔用空間時效能會降低 clustering 越來越大

Q5. What is "secondary clustering," and how is it different from primary clustering?

A5: 和 primary clustering 不同 secondary clustering 會繼續排序
即 clustering 不會和 primary clustering 一樣越來越大.

Q6. Briefly explain the difference between:

- Open addressing
- Separate chaining

A6: Open addressing 開放定址法. \Rightarrow 存在同一 hash table. 碰撞了就找下一個空位.
separate chaining 鏈結法 \Rightarrow 用 linked list 方式

array + linked list.

Part B: Hash Function Calculation (Collision & Pattern Observation)

Show your steps clearly.

Hash Function 1 — Division Method

$$h_1(k) = k \bmod 10$$

Hash Function 2 — Folding Method

Split key into two-digit chunks and sum the chunks.

$$h_2(k) = (\text{sum of 2-digit groups}) \bmod 11$$

Example:

Key = 8429 \rightarrow groups: 84 + 29 \rightarrow 113 \rightarrow 113 mod 11 = 3

Q7. (Compute using Hash Function 1)

Given keys: 27, 37, 47, 57, 67

Compute their hash values using:

A7: $27 \bmod 10 = 7$ $37 \bmod 10 = 7$ $47 \bmod 10 = 7$ $57 \bmod 10 = 7$ $67 \bmod 10 = 7$ $h_1(k) = k \bmod 10$
A: 7, 7, 7, 7, 7.

Q8. (Identify collision pattern)

From your results in Q1:

- What pattern do you observe?
- Explain why these keys collide.

A8: 每個數對10取餘後都是7 \rightarrow hash values 都是7
hash functions 太簡單了, 無法有效分配這些 keys.

Q9. (Compute using Hash Function 2)

Compute $h_2(k)$ for: 1234, 9217, 4519, 9902

A9: $1234 = 46$ $46 \bmod 11 = 2$
 $9217 = 109$ $109 \bmod 11 = 10$
 $4519 = 64$ $64 \bmod 11 = 9$
 $9902 = 101$ $101 \bmod 11 = 2$
A: 2, 10, 9, 2.

Q10. (Compare distribution)

- Which hash function (h_1 or h_2) produced more collisions for the input set?
- Which seems to spread keys more evenly?
- Provide 1–2 sentences of explanation.

A10: 1. h_1 , h_1 產生出的 hash values 都是7
2. h_2 , h_2 產生出的 hash values 仍有碰撞的情況
但因 hash functions 設計得較完善, 所以 keys 更加分散.