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

Quiz III: Hash Table

November 25, 2025, 16:40~17:00

Student ID: 113153 Student Name: 蕭承萱 100

+50

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 被分配給同個 index

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

A2: ◦ [A] → "Apple" → "Analyze"

↓

bucket. (儲存位置/儲存桶)

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

A3: load factor 負載因子.

⇒ $\frac{\text{elements}}{\text{total buckets}}$ ⇒ 比例越高 ⇒ collision 越多.

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

A4: primary clustering 初次聚集 linear probing.

⇒ open addressing 使有一大串連續方格用空間時，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 second clustering 會繼續排序

和 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. ⇒ 存在同 hash table (array) 碰撞了就找下一個空位.

Separate chaining 連結法 ⇒ 用 linked list 方式放在 bucket 中

↓

array

+

linked list.

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

Show your steps clearly.

Hash Function 1 — Division Method

$$h_1(k) = k mod 10$$

Hash Function 2 — Folding Method

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

$$h_2(k) = (sum\ of\ 2-digit\ groups) mod 11$$

Example:

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

Q7. (Compute using Hash Function 1)

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

Compute their hash values using:

$h_1(k) = k mod 10$

A7: 27 mod 10 = 7　57 mod 10 = 7

37 mod 10 = 7　67 mod 10 = 7.

47 mod 10 = 7

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 ⇒ 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: 12+34 = 46　46 mod 11 = 2

92+17 = 109　109 mod 11 = 10

45+19 = 64　64 mod 11 = 9

99+2 = 101　101 mod 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 產生過的 hash values 都是 7

2. h_2, h_2 產生過的 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.