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

Quiz III: Hash Table

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

Student ID: 113528 Student Name: 戴光彥 70

+10

Part A: Hash Table Definitions (Conceptual Understanding)

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

A1: 不同 key 被分配到同個 index 時，即是 collision

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

A2: 存放一筆或多筆資料的容器

+5

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

A3: $\alpha = n/m$；n=元素之數量；m=表格大小

更高的 α → 更多碰撞 → 更長的探測鏈

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

A4: 在發生 collision 時，會選擇往下個 index 去尋找空間儲存．線性探測，形成長而連續的填充槽

+5

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

A5: 具有相同 Hash 值的 key 遵循同個探測序列；不連續但相同的模式．

Q6. Briefly explain the difference between:

- Open addressing
- Separate chaining

A6: Separate chaining 在遇到 collision 時，是運用 linked lists，去指向被分配在同個 index 的元素．

Open addressing 則是存在同個 array，去往後尋找未被使用的空間去存放。

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.

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

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:

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

A7:

27 => 27 mod 10 = 7, 37 => 37 mod 10 = 7, 47 => 47 mod 10 = 7

57 => 57 mod 10 = 7, 67 => 67 mod 10 = 7

Q8. (Identify collision pattern)

From your results in Q1:

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

A8: 所有 key 經過 Hash function 1 處理後，都得出了相同結果，發生了 collision，因為每個 key 的個位數皆為 7，單單 mod 10 的結果因此相同。

Q9. (Compute using Hash Function 2)

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

A9:

1234 => (12+34) mod 11 = 2

9217 => (92+17) mod 11 = 10

4519 => (45+19) mod 11 = 9

9902 => (99+02) mod 11 = 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: h_1 發生較多 collision，h_2 則使 keys 分散的較平均。h_1 因只進行 mod 10，當 keys 的尾數相同，不論數字多大（507, 897, 27, 1047），都會發碰撞，而 h_2 在面對同問題時，把數字拆開相加，能有效打散數字