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

Quiz III: Hash Table

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

85

Student ID: 113709 Student Name: 李啓嘗

Part A: Hash Table Definitions (Conceptual Understanding)

+35

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

A1: 如果key指向同一個index會發生碰撞

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

A2: 儲存資料的地方

+5

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

A3: n/m , n是有幾個key被儲存 m是index有幾個
n跟m越接近越容易碰撞

+5

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

A4: 會聚集, linear probing (h(k) + i) mod 10

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

A5: 會因為key指向同一個index,間隔越跳越遠
主要發生在 (h(k) + c1 + c2i^2) mod 10

+5

Q6. Briefly explain the difference between:

- Open addressing
- Separate chaining

A6: open addressing: array, 儲存資料只有固定空間
separate chaining: linked list, 可以一直往下接

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 % 10 => 7
37 => 37 % 10 => 7
47 => 47 % 10 => 7
57 => 57 % 10 => 7
67 => 67 % 10 => 7

Q8. (Identify collision pattern)

From your results in Q7:

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

A8: 全部 index 都是 7，因為個尾數是 7 且是除以 10 的餘數

Q9. (Compute using Hash Function 2)

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

A9:
1234 => 12+34 => 46 % 11 => 2
9217 => 92+17 => 109 % 11 => 10
4519 => 45+19 => 64 % 11 => 9
9902 => 99+2 => 101 % 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 碰撞較多，h_2 把 key 分比較開較隨機
因為 h_1 mod 的數是 10，所以碰到個位數相同時，會發生碰撞，h_2 mod 的數是 11 是質數而且有先對 key 做處理