Course: Data Structures (CSE CS203A, 114-1)
Quiz III: Hash Table
November 25, 2025, 16:40~17:00

72.5

Student ID: 113332✓ Student Name: 林晉逵

Part A: Hash Table Definitions (Conceptual Understanding)

+25

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

A1: 不同的 key 經 hash function 得到相同的 index (hash value)，造成碰撞

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

A2: key $\xrightarrow{\text{hash function}}$ index (hash value) 映射出 value (存放 value 的地方就是 bucket)

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

A3: $\frac{\text{elements 的數量}}{\text{tablesize}}$，∝越大 element 分部的越集中，反之亦然

+5

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

A4: 不同key經hash function得到大量相同的index (hash value)，導致elements大量堆積在同一處。 linear probing

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

A5: 不同的key經hash function得到大量相同的index，導致elements大量堆積在同一處。

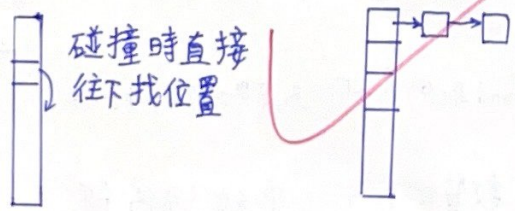
線性遞增 $(\text{hash}(x)+ci+ci^2)$

hash function 的不同

Q6. Briefly explain the difference between:

- Open addressing
- Separate chaining

A6:



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, 37, 47, 57, 67

↓ ↓ ↓ ↓ ↓

hash values: 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: ① 全部的keys都有相同的hash value ② 因為全部的keys尾數都是7，且mod又取10，造成了所有的keys都碰撞在一起

Q9. (Compute using Hash Function 2)

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

A9: 1234 → (12+34) mod 11 = 1

9217 → (92+17) mod 11 = 9

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? 分散
- Provide 1–2 sentences of explanation.

A10: ① h_1 (27, 37, 47, 57, 67) → 碰撞

h_2 (9217, 4519) → 碰撞

② h_2

③ (1) mod 取的數不同，取質數分佈會較分散

(2) 數的拆分，也使得keys較分散