Student ID: 113417   Student Name: 魏允鳩   100

Part A: Hash Table Definitions (Conceptual Understanding)

50

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

A1: A collision happens when two different keys are hashed to the same index in the hash table.

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

A2: A bucket is a storage location (or slot) in the hash table where a key – value pair is placed.

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

A3: ① The "load factor (α)" = (number of stored items) / (table size)
② A higher α means more collision, so search and insert become slower.

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

A4: ① Primary clustering is when long blocks of consecutive filled slots form, causing long probe sequences.
② It occurs in linear probing.

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

A5: ① Secondary clustering : Keys with the same hash index follow the same probe pattern.
② Happens in quadratic probing, and clusters are smaller than primary clusting.

Q6. Briefly explain the difference between:
- Open addressing
- Separate chaining

A6: Open addressing : Store everything inside the table, resolve collision by probing.

Separate chaining : Each bucket has a list that can store multiple items.

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 mod 10 = 7   57 mod 10 = 7

37 mod 10 = 7   67 mod 10 = 7

47 mod 10 = 7

Q8. (Identify collision pattern)

From your results in Q1:

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

A8: ① All keys end with digit 7,

② h_1 only uses the last digit, so keys with the same ending collide.

Q9. (Compute using Hash Function 2)

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

A9:

1234 → 12+34 = 46 → 46 mod 11 = 2

9217 → 92+17 = 109 → 109 mod 11 = 10

4519 → 45+19 = 64 → 64 mod 11 = 9

9902 → 99+2 = 101 → 101 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: ① h_1 produced more collision.

② h_2 spread keys more evenly.

③ h_1 only depends on the last digit, so many keys collide.

h_2 uses more of the key's digits and spreads keys better.