PW SKILLS

$hf(v) = v \% m$

Collision Resolution Technique

key	value
0	
1	
2	
3	
4	24
5	
6	76
7	
8	
9	39

hash function → $v \% m$

24, 59, 76, 69, 44, 65

69 → 39 (Collision)

$m = 10$

Advantage

We can store infinite num of value

Limitation

1) search → $O(n)$

2) Extra space outside of hash-table (Space is available inside hash table) ①

1) chaining → Linked List

2) Open addressing → 1) Linear Probing ②, 2) Quadratic Probing ③, 3) Double Hashing ③, 4) ④

$m = 10$

value = 57, 67, 97, 45, 62, 107

Hash Table

key	value
0	–
1	–
2	5000
3	–
4	–
5	4000
6	–
7	1000
8	–
9	–

chaining

5000: [62 | null]

4000: [45 | null]

1000: [57 | 2000] → 2000: [67 | 3000] → 3000: [97 | 6000] → 6000: [107 | null]

$57 \% 10 = 7$, $97 \% 10 = 7$, $67 \% 10 = 7$, $45 \% 10 = 5$, $62 \% 10 = 2$, $107 \% 10 = 7$

Linked List / chain

Linear Probing

65

Values = 50, 75, 99, 20, 35, 88, 45, 23, 55, 67

$LP(v,i) = (hf(v) + i) \% m$

- m → size of hash table
- i → # collisions

$hf(v) = v \% m$

Hash Table

key	value
0	50
1	20
2	55
3	23
4	67
5	75
6	35
7	45
8	88
9	99

$LP(20,1) = (0+1) \% 10 = 1$

$LP(35,1) = (5+1) \% 10 = 6$

$LP(45,1) = (5+1) \% 10 = 6$

$LP(45,2) = (5+2) \% 10 = 7$

$LP(55,1) = (5+1) \% 10 = 6$

$LP(55,2) = (5+2) \% 10 = 7$

$LP(55,3) = (5+3) \% 10 = 8$

$LP(55,4) = (5+4) \% 10 = 9$

$LP(55,5) = (5+5) \% 10 = 0$

$LP(55,6) = (5+6) \% 10 = 1$

$LP(55,7) = (5+7) \% 10 = 2$

Search → time complexity

Worst case → $O(n)$ (55)

Best case → $O(1)$ (75)

55 → $hf(55) = 5$

65 → $hf(65) = 5$

→ Primary clustering

Quadratic Probing

Values = 50, 75, 99, 20, 35, 88, 45, 23, 55, 67

$c_1 = c_2 = 1$

$$Q\,P(v,i) = (hf(v) + c_1 i + c_2 i^2) \% m$$

(m: size of hash table; i: # collisions)

Hash Table

key	value
0	50
1	45
2	20
3	23
4	
5	75
6	
7	35
8	88
9	99

20 → 50

55 → 45

$Q\,P(20,1) = (0+1+1)\%10$

$Q\,P(35,1) = (5+1+1)\%10$

$= 7$

55 → 75

$Q\,P(45,1) = (5+1+1)\%10$

$= 7$

55 → 35

$Q\,P(45,2) = (5+2+4)\%10$

$= 11\%10 = 1$

$Q\,P(55,3) = (5+3+9)\%10$

$= 7$

$Q\,P(55,4) = (5+4+16)\%10$

$= 5$

$Q\,P(55,5) = (5+5+25)\%10$

$= 5$

Limitation → Loss of data (55), Secondary clustering,
search → O(m)

Double Hashing

value = 50, 75, 99, 20, 35, 88, 45, 23, 55, 67

(# collisions)

$$DH(v,i) = (hf_1(v) + i\,hf_2(v)) \% m$$

$$hf_2(v) = 1 + v \bmod (m-2)$$

0	50
1	
2	
3	
4	
5	75
6	
7	
8	
9	99

20

$hf_2(20) = 1 + 20 \% 8$

$= 1 + 4$

$= 5$

$hf_1(20) = 0$

$DH(20,1) = (0+5) \% 10$

$= 5$

20

$DH(20,2) = (0 + 2*5) \% 10$

$= 0$

$DH(20,3) = (0 + 3*5) \% 10$

$= 5$

↳ Loss of Data

↳ Search → O(m)