

Hw6_22000123_NguyenDucSi_Baitap3

Nguyễn Đức Sĩ

November 2024

0.0.1 Câu 1

Consider a hash table of size $m = 10000$, and the hash function $h(K) = \lfloor m(KA \bmod 1) \rfloor$ for $A = \frac{\sqrt{5}-1}{2}$. The key 123456 is mapped to location ____.

$$\begin{aligned} \text{Có } KA \bmod 1 &= 123456 * \frac{\sqrt{5}-1}{2} = 0.00412 \\ \Rightarrow \lfloor m(KA \bmod 1) \rfloor &= \lfloor 41.2 \rfloor = 41 \end{aligned}$$

0.0.2 Câu 2

Consider a 13 element hash table for which $f(\text{key}) = \text{key} \bmod 13$ is used with integer keys. Assuming linear probing is used for collision resolution, at which location would the key 103 be inserted, if the keys 661, 182, 24 and 103 are inserted in that order?

$$\text{Có } 661 \bmod 13 = 11, 182 \bmod 13 = 0, 24 \bmod 13 = 11$$

182											661	24
0	1	2	3	4	5	6	7	8	9	10	11	12

$$\text{Có } 103 \bmod 13 = 12 \Rightarrow 103 \text{ được chèn vào ô 1}$$

0.0.3 Câu 3

Consider a hash table of size $m = 100$ and the hash function $h(k) = \text{floor}(m(kA \bmod 1))$ for $A = \frac{\sqrt{5}-1}{2} = 0.618033$. Compute the location to which the key $k = 123456$ is placed in hash table.

$$\text{Có } h(k) = \text{floor}(m(kA \bmod 1)) = \text{floor}(100(123456 * 0.618033 \bmod 1) = 88$$

$$\Rightarrow \text{chèn vào ô 88}$$

0.0.4 Câu 4

A hash table with 10 buckets with one slot per bucket is depicted here. The symbols, S1 to S7 are initially entered using a hashing function with linear probing. The maximum number of comparisons needed in searching an item that is not present is

0	S7
1	S1
2	
3	S4
4	S2
5	
6	S5
7	
8	S6
9	S3

Ta có từ ô số 8 đến 9 và quay lại từ ô 0 đến 1 là 4 ô liên tiếp, do vậy nếu điền vào ô 8, ta cần phải tìm tối đa là 5 ô để chèn được số đó.

0.0.5 Câu 5

A hash function h defined $h(\text{key}) = \text{key} \bmod 7$, with linear probing, is used to insert the keys 44, 45, 79, 55, 91, 18, 63 into a table indexed from 0 to 6. What will be the location of key 18 ?

Có $h(44) = 44 \bmod 7 = 2$, $h(45) = 45 \bmod 7 = 3$, $h(79) = 79 \bmod 7 = 2$, $h(55) = 55 \bmod 7 = 6$, $h(91) = 91 \bmod 7 = 0$, $h(18) = 18 \bmod 7 = 4$, $h(63) = 63 \bmod 7 = 0 \Rightarrow$ thứ tự điền vào là : 91, 63, 44, 45, 79, 18, 55, ...
 \Rightarrow Điền vào ô 5

0.0.6 Câu 6

Consider a double hashing scheme in which the primary hash function is $h_1(k) = k \bmod 23$, and the secondary hash function is $h_2(k) = 1 + (k \bmod 19)$. Assume that the table size is 23. Then the address returned by probe 1 in the probe sequence (assume that the probe sequence begins at probe 0) for key value $k = 90$ is _____.

Kích thước bảng = 23, giá trị khóa $k = 90$, Và, đầu dò $i = 1$ (đầu dò 1 trong chuỗi đầu dò).

Ta có :

$$h_1(k) = k \bmod 23 = 90 \bmod 23 = 21$$

Và,

$$h_2(k) = 1 + (k \bmod 19) = 1 + (90 \bmod 19) = 15$$

Sử dụng băm kép, ta có:

$$(h_1(k) + i \cdot h_2(k)) \bmod \text{size} = (21 + 15) \bmod 23 = 36 \bmod 23 = 13$$

0.0.7 Câu 7

How many different insertion sequences of the key values using the hash function $h(k) = k \bmod 10$ and linear probing will result in the hash table shown below?

0	
1	
2	42
3	23
4	34
5	52
6	46
7	33
8	
9	

Ta có 42, 23, 34 phải xuất hiện trước 52 và 46 phải xuất hiện trước 33. Tức với 42, 23, 34 xuất hiện trước, ta có 3! cách chọn. Với 46, ta có thể chọn xuất hiện ở bất kỳ đâu trước số 33, tức ta có 5 cách, và còn lại 52 và 33 sẽ xuất hiện cuối cùng. \Rightarrow có $3! \cdot 5 = 30$ cách.

0.0.8 Câu 8

Which of the following hash functions is most likely to cause clustering in a hash table?

1. $h(k) = k \bmod m$
2. $h(k) = \lfloor m \cdot (kA \bmod 1) \rfloor$
3. $h(k) = k$
4. $h(k) = \left(\frac{k}{m} + k \cdot m\right) + k \bmod m$

Ta có phép chia lấy dư luôn dễ gây ra đùng hơn khi các số cho cùng một kết quả dư, các hàm băm khác phức tạp hơn sẽ giúp cho sự phân bố của các số đều hơn. \Rightarrow chọn 1

0.0.9 Câu 9

Consider a hash table with 100 slots. Collisions are resolved using chaining. Assuming simple uniform hashing, what is the probability that the first 3 slots

are unfilled after the first 3 insertions?

Ta có 3 ô đầu không được điền sau 3 lần chèn đầu tiên thì xác suất sẽ là :

$$\frac{97 \times 97 \times 97}{100^3}$$

0.0.10 Câu 10

Which one of the following hash functions on integers will distribute keys most uniformly over 10 buckets numbered 0 to 9 for i ranging from 0 to 2020?

Xét chữ số cuối:

Số	Bình phương	Lập phương
0	0	0
1	1	1
2	4	8
3	9	7
4	6	4
5	5	5
6	6	6
7	9	3
8	4	2
9	1	9

Ta có thể thấy rằng với phép chia 10 ta chỉ quan tâm đến số cuối cùng, và của Lập phương thì phân bố của lập phương trải đều hơn số với bình phương. Với $(12 * i)$ thì có ước chung là 2, nên phân bố cũng không tốt bằng, và $11 * i^2$ thì số cuối cũng phân bố tương tự khi chỉ bình phương.

=> chọn $i^3 \bmod 10$.

0.1 Câu 11

Consider a hash function that distributes keys uniformly. The hash table size is 20. After hashing of how many keys will the probability that any new key hashed collides with an existing one exceed 0.5.

Đối với mỗi mục nhập, xác suất va chạm là $1/20$ Sau khi chèn các giá trị x, xác suất là $1/2 \Rightarrow (1/20).x = 1/2 \Rightarrow X=10$

0.2 Câu 12

What is the probability of a collision when hashing n keys into a hash table of size m, assuming that the hash function produces a uniform random distribution?

Khi số lượng mục tăng lên, xác suất va chạm cũng tăng lên, khi kích thước của bảng băm tăng lên, xác suất giảm xuống. Do đó, xác suất va chạm có thể được ước tính là $O(n/m)$.

0.3 Câu 13

Insert the characters of the string K R P C S N Y T J M into a hash table of size 10. Use the hash function $h(x) = (ord(x) - ord("A") + 1) \bmod 10$. If linear probing is used to resolve collisions, then the following insertion causes collision

Ta có :

K sẽ được chèn vào chỉ mục: $(11-1+1) \bmod 10 = 1$

R tại chỉ mục: $(18-1+1) \bmod 10 = 8$

P tại chỉ mục: $(16-1+1) \bmod 10 = 6$

C tại chỉ mục: $(3-1+1) \bmod 10 = 3$

S tại chỉ mục: $(19-1+1) \bmod 10 = 9$

N tại chỉ mục: $(14-1+1) \bmod 10 = 4$

Y tại chỉ mục: $(25-1+1) \bmod 10 = 5$

T tại chỉ mục: $(20-1+1) \bmod 10 = 0$

J tại chỉ mục: $(10-1+1) \bmod 10 = 0 \Rightarrow$ xảy ra va chạm đầu tiên.

M tại chỉ số: $(13-1+1) \bmod 10 = 3 \Rightarrow$ xảy ra va chạm thứ hai. Chỉ có J và M gây ra va chạm

0.4 Câu 14

What is the space used by programs when we use Hashmap in our program?

Không gian được sử dụng bởi hashmap trong chương trình là $O(N)$,

0.5 Câu 15

Which searching technique takes $O(1)$ time complexity for searching the data?

Băm có độ phức tạp $O(1)$ để tìm kiếm dữ liệu.