Hashing PDF questions:

①

0	37
1	91
2	
3	54
4	52
5	
6	55
7	
8	56
9	53

Linear probing:

$H(x,i) = (h(x) + i) \% SIZE$

Let's say key y was searched in the table and none of $s_1 - s_4$ is y

$y = 20$:

$H(20,0) = (\underbrace{(20 \% 10)}_{0} + \underbrace{0}_{}) \% SIZE$
$= (0) \% 10$
$= 0$

But $S4 \neq y$

Next, collision increase
$H(20,1) = (0+1) \% 10$
$= 1$

$S1 \neq 20$

Next,
$H(20,2) = (0+2) \% 10$
$= 2$

Address 2 is empty.
Now, we definitely know that 20 is not present in the hash table.

To answer the question, we need to find the longest sequence of filled addresses followed by an empty address

$y = 28$

$\begin{cases} i=0 \Rightarrow H(x) = 8 \\ i=1 \Rightarrow H(x) = 9 \\ i=2 \Rightarrow H(x) = 0 \\ i=3 \Rightarrow H(x) = 1 \\ i=4 \Rightarrow H(x) = 2 \Rightarrow empty \end{cases}$

$\longrightarrow 5$ comparisons //

②

Ⓐ In worst case, in direct chaining, all keys are in one LL.
$\Rightarrow$ So, TC = $O(n)$

In open addressing, in worst case, we go through the entire table SIZE-1 times. i.e.
we start at addr. x and circle the entire table back to $x-1$

$\Rightarrow$ So, TC = $O(SIZE)$

That is not necessarily true since we don't know if $n < SIZE$

Ⓑ We use direct chaining when no. of keys > hash table size.
So, space used is def. not less.

Ⓒ Deletion involves searching back to A.

So, option D, none of the above

TC is same for both $\Rightarrow O(n)$

✗ Direct chaining requires more space.

Actually if we do knowchaining that that n > SIZE use that is when we use direct chaining.

DON'T LIKE THIS REASONING!

Open addressing is worse at deletion since we need to recompute for every key

③ i & ii only

④

ⓘ hash func may result in collision ✓

ⓘⓘ i/p is fixed length ✗
Not necessarily
I/Ps: 1439
123

ⓘⓘⓘ i/p is of arbitrary length ✓
but the fixed length code part $\rightarrow$ not sure

I & III

$\rightarrow$ Size of hash address is same so fixed length code part is true

⑤

Probability of 1 slot being filled

Prob. of not getting any of n slots
$= \frac{n}{60}$

Uniform hashing $\Rightarrow$ Considering this to mean that only 1 insertion per slot

$\Rightarrow$ 1st insertion $= \frac{(100-1)}{100} = \frac{99}{100}$ ← in slot not filled

$\Rightarrow$ 2nd $\frac{(99-1)}{100} = \frac{98}{100}$ ← 1 slot filled, 1 slot not filling

$\Rightarrow$ 3rd " $\frac{98-1}{100} = \frac{97}{100}$

$\Rightarrow \frac{99}{100} \times \frac{98}{100} \times \frac{97}{100}$

✗

⑥ $\alpha = \frac{n}{m} = \frac{30000}{250} = 80$

α needs to be greater than 1.

⑦

Double hashing $= \left(h_1(x) + i\, h_2(x)\right) \% SIZE$

$i = 1$
$h_1(x) = x \% 23$
$h_2(x) = 1 + x \% 19$

$(90 \% 23 + (1 + 90 \% 19)) \% 23$

$\Rightarrow (21 + (1+14)) \% 23$

$\Rightarrow 13$

23) 90 (3
   -69
   ___
    21

19) 90 (4
    76
    ___
    14

21
14
___
35

23) 35 (1
    23
    ___
    12

23) 90 (
   69
   ___
   23 ×3
     69

14) 90 (
    76
    ___
    39 ×4
    76

Forgot the +1 in h2
Ended up getting the 12 instead of 13. This was a numerical question.

⑧ $h(x) = (3x+4) \% 7$

$h(1) \Rightarrow (3(1)+4) \% 7 \Rightarrow 0$
$h(3) \Rightarrow (9+4) \% 7 \Rightarrow 13 \% 7 \Rightarrow 6$
$h(8) \Rightarrow (24+4) \% 7 \Rightarrow 28 \% 7 \Rightarrow 0$
$h(10) \Rightarrow (30+4) \% 7 \Rightarrow 34 \% 7 \Rightarrow 6$

1,8	0					3,10
0	1	2	3	4	5	6

Unsure what closed hashing is. Assuming it be linear probing.

for 8:
$i = 1$
$H(x) = (h(x)+i) \% SIZE$
$= (0+1) \% 7$
$= 1$
$H(8) \Rightarrow 1$

for 10:
$i = 1$:
$H(x) = (h(x)+i) \% SIZE$
$= (6+1) \% 7$
$= 0$

$i = 2$:
$H(10) = (6+2) \% SIZE$
$= 8 \% 7$
$= 1$

$i = 3$:
$H(10) = (6+3) \% 7$
$= 9 \% 7$
$= 2$

$H(10) \Rightarrow 2$

1	8	10				3
0	1	2	3	4	5	6

⑨ $n = 20$

If by uniformly they mean that every key can be stored in the table

$\frac{Places}{key}$

$\Rightarrow 20$

⑩

i^2:
0 1 4 9 6 5 6 9 4 1

6 unique values

i^3:
0 1 8 7 4 5 6 3 2 9
10 unique values

$11 * i^2$:
0 1 4 9 6 5 6 9 4 1
6 unique values

$12 * i$:
0 2 4 6 8 0 2 4 6 8
5 unique values

⑪ Question incomplete

what is the expected no. of probes in successful search?

$\alpha = \frac{n}{m} = \frac{9800}{10000} = 0.98$ | $m = 10000$
$n = 9800$

$\alpha > 1$ means not enough space for keys
i.e. keys > slots

$\Rightarrow \frac{n}{m} > 1 \Rightarrow$ $m \rightarrow slots$
$n \rightarrow keys$

$\frac{1}{\alpha} * \ln\left(\frac{1}{1-\alpha}\right)$

$\Rightarrow \frac{1}{0.98} * \ln\left(\frac{1}{0.02}\right)$

$\Rightarrow \frac{100}{98} * \ln\left(\frac{100}{2}\right)$

$\Rightarrow \frac{100}{98} * \ln(50)$

$\Rightarrow 3.99$ //

So, C

Prob. of 1st insertion not being in 0,1 & 2 addresses $\Rightarrow 97/100$

2nd insertion $\Rightarrow$ also $97/100$

3rd insertion $\Rightarrow 97/100$

$\Rightarrow \frac{(97)^3}{(100)^3}$

$\rightarrow$ Why? B/c this is chaining Every insertion could be on the same address