Information

Dodo Hashing

Dodo hashing is an implementation of a symbol table invented by the species *Raphus cucullatus*, who based their society's IT infrastructure on it.

The idea is as follows. We'll assume all the keys are nonnegative integers just to keep it simple, and the hash value is simple modular hashing. To be precise, to map the key k to an array index between 0 and $M-1$ we compute $h(k) = k \mod M$. In pseudo code,

```
function hash(x: int, M: int) -> int:
    return x % M
```

The idea is simple: store k in a table of `int` of length `M` at `T[hash(k, M)]`. Dodo society had not heard of collision strategies such as chained hashing, open addressing, or other avian-themed ideas, so their strategy was

> whenever a collosion occurs, increase `M` by one, initialise a new table with `M` entries, initially all set to `null` or `None`, and move all the keys into a new table

For instance, `M=3` and the keys stored are $10, 21$. Then the table looks like this (not showing the `null` / `None` entries):

```
 0 | 1 | 2
---+---+---+
 21| 10|
```

Now `6` is added to the table, which should go to `hash(6,3)=0`, but that position is already taken. `M` is increased to `4`, and everthing is moved into a new table:

```
 0 | 1 | 2 | 3
---+---+---+---
   | 21| 10|
```

Alas, `6` still doesn't fit (it wants to go to position `2`, which is taken), so `M` is increased further. This continues until all keys have found their own cell.

Question 11

Not yet answered

Marked out of 1.00

I pick a random number k and insert it into an empty Dodo table of size M. Then I insert the integer $M + k$. What is the total running time of both operations?

- a. $O(N)$, linear in the number of keys
- b. $O(1)$
- c. $O(M)$, linear in the table size
- d. $O(\log(N + M))$, logarithmic in the total size

Clear my choice

Question 12

Answer saved

Marked out of 1.00

Assume two keys, picked independently at random, are inserted into an empty Dodo-hash data structure. What is the expected total running time for these two operations?

- a. Infinite
- b. Linear
- c. Logarithmic
- d. Constant

Clear my choice

Question 13

Answer saved

Marked out of 1.00

Assume a Dodo hash table contains $M/2$ keys. A new random key is inserted. What is the probability of a collision?

- a. 0
- b. 1
- c. $\frac{1}{2}$
- d. $M/2$

Clear my choice

Question 14

Answer saved

Marked out of 1.00

What is the worst-case running time of inserting a single element into a Dodo table of size M holding N elements?

- a. $O(N)$
- b. $O(M)$
- c. $O(NM)$
- d. $O(1)$

Clear my choice

Question 15

Answer saved

Marked out of 1.00

(Hard) Assume I want to store N random keys in my Dodo-hashing-based symbol table. How large does the table have to be so that the probability that none of the insertions have a collision is $\frac{1}{2}$?

- a. quadratic in N because of the birthday problem
- b. linearithmic in N because of the coupon collector's problem
- c. exponential in N because you insert N elements and there is a $\frac{1}{2}$ collision probability for each
- d. linear in N because then M/N is a constant

Clear my choice