Take-Home Final Exam

due Monday, March 14, 2022

Note: if you are in CIS413, then do four of the following five problems. If you are in CIS513, then you need to do all five problems.

1. Instead of array doubling, here we will examine array tripling. While inserting elements into an array (from left to right) and there is a free space in the desired location $i$, we write the new element there (cost 1). On the other hand, if the array is full, we will create a new array that is triple the size of the original one and copy all elements from the old array into the new one (and then perform the write). We are charging only for the copy (no allocation-initialization steps).

So in this case, if the array is of size $N$ (where $N = 3^k$ for some $k$) and there is an insertion at position $i = N + 1$, we create a new array of size $3N$. Then we copy the $N$ elements in the old array to the new one: the total expansion cost is $N$. Next we write the new element at position $i = N + 1$ for an additional cost of 1 (overall actual cost is $N + 1$).

Using the potential function $\Phi = \frac{1}{2}(2f - e)$, where $f$ is the number of occupied (full) locations in the array and $e$ is the number of unused (empty) locations, derive the amortized cost of an INSERT operation.

Note: If you want to use the measures size and cap (where size is $f$ and cap is $f + e$), then $\Phi = \frac{1}{2}(3\text{size} - \text{cap})$. Also, the amortized cost of an operation does not need to be an integer.

2. (exercise 13, from chap 11 of Er)
The Department of CS at SPU has a flexible curriculum with a complex set of graduation requirements. The department offers $n$ different courses and there are $m$ different requirements. Each requirement specifies a subset of the $n$ courses and the number of courses that must be taken from that subset. The subsets for different requirements may overlap, but each course can be used to satisfy at most one requirement. For example, suppose there are $n = 5$ courses $A, B, C, D, E$ and $m = 2$ graduation requirements:

- you must take two courses from the subset $\{A, B, C\}$
- you must take two courses from the subset $\{C, D, E\}$

Then a student who has taken courses $B, C, D$ cannot graduate, but a student who has taken either $A, B, C, D$ or $B, C, D, E$ can graduate.

Describe and analyze an algorithm to determine whether a given student can graduate. The input to your algorithm is the list of $m$ requirements (each specifying a subset of the $n$ courses and the number of courses that must be taken from each subset) and the list of courses that the student has taken.

3. (exercise 19, part (a) from chap 11 of Er, reworded)
You are given (update) a directed graph $G = (V, E)$ with start and finish nodes $s, t \in V$. The nodes represent locations and edges are roads between the locations along which people
can walk. Each edge/road of $G$ is gated, allowing at most one person to walk through per hour: for example one person can follow that edge between 1:00 and 1:05, and then another between 2:00 and 2:05.

Given an integer $h$, describe an algorithm and give the time to determine the maximum number of people who can walk from $s$ to $t$ in $h$ hours. (The hint from Er indicates that you need to build a new graph and the time bound will include $h$.)

4. Suppose we have three sets $S_0, S_1, S_3 \subseteq \mathcal{U}$ represented by Bloom filters $B_0, B_1, B_3$ respectively. All sets have $n$ elements and all filters are of size $m$ and were constructed using the same set of hash functions $h_1, h_2, \ldots, h_k : \mathcal{U} \to [m]$. (The parameters $n, m, k$ are the same as used in the Er text.)

(a) Let $B$ be the bit-wise AND of the vectors $B_0$ and $B_1$, which represents the set $S = S_0 \cap S_1$. What is the false positive rate for a query to $B$?

(b) Let $B$ be the bit-wise OR of the vectors $B_0$ and $B_1$, which represents the set $S = S_0 \cup S_1$. What is the false positive rate for a query to $B$?

(c) Suppose that you manage to construct a set $S \subseteq \mathcal{U}$ of size $n$ such that by storing it in a Bloom filter $B$ using the hash functions above you obtain $B = B_3$ (that is, the vectors have the same pattern of zeroes and ones). You want to say you have a “copy” of $S_3$. What is the probability that $S = S_3$?

5. Consider an array containing the following 20 values:

$$3, 5, 2, 10, 9, 2, 3, 2, 8, 4, 5, 1, 6, 12, 7, 9, 8, 8, 4, 1$$

(a) Draw the segment tree allowing range queries on these values.

(b) Draw the corresponding Binary Indexed Tree (= Fenwick tree) for the same.