York University
EECS 4101/5101, Winter 2023
Assignment 5

Due Date: April 11th, at 23:59

It's your road, and yours alone.
Others may walk it with you, but no one can walk it for you ...

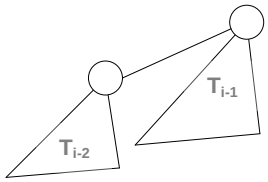
Rumi

All problems are written problems; submit your solutions electronically **only via Crowdmark**. You are welcome to discuss the general idea of the problems with other students. However, you must write your answers individually and mention your peers (with whom you discussed the problems) in your solution. There is also a bonus question, which is harder than other questions. It is recommended to approach this question only if you have completed other questions. Please refer to the course webpage for guidelines on academic integrity.

Problem 1 Binomial Tree Variant [4 + 6 + 6 = 16 marks]

In the class, we learned that a binomial tree of order i can be formed by taking two copies of a binomial tree of order $i-1$ and letting the root of one tree have the other tree as its first child. Consider the following definition of **Fibomial** trees:

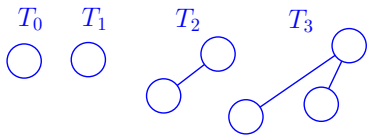
- A Fibomial tree of order 0 is a single node and a Fibomial tree of order 1 is also a single node.
- A Fibomial tree of order i is formed by taking a Fibomial tree of order $i-1$ and a Fibomial tree of order $i-2$ letting the root of the first tree (of order $i-1$) have the other tree as its first child (see the figure below).



a) Indicate whether the following statement is correct or not. Provide a justification of your answer.

Let T_k denote the Fibomial tree of order k. The children of the root of T_k are the Fibomial trees $T_{k-2}, T_{k-3}, \ldots, T_2, T_1, T_0, T_0$ (more precisely, there is exactly one copy of each T_i for $i \in \{k-2, k-3, \ldots, 1\}$ and two copies of T_0 as a children of T_k).

Answer: The statement is incorrect. For example, consider T_3 in the example below; if the statement was correct, there should have been 3 leaves (one for the tree of order 1 and two trees of order 0).



b) Indicate how many nodes exist in a Fibomial tree of order k. You should provide a recursive formula and relate it to a known series.

Answer:
Let $N(T_k)$ be the number of nodes in a Fibomial tree of order k. Then we have:

$$N(T_k) = \begin{cases} 1 & k = 0 \\ N(T_{k-1}) + N(T_{k-2}) & k > 0 \end{cases}$$

which gives $N(T_k) = F(k-1)$, the $k-1$'th Fibonacci number. We saw in the class that how to find the exact value of Fibonacci number, as well as a lower bound for it.

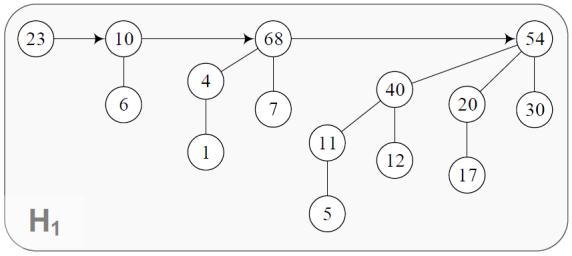
c) Indicate what the height of a Fibomial tree of order k is. Again, you should provide a recursive formula and solve it. **Answer:**
Let $h(T_k)$ be the height of a Fibomial tree of order k. Then we have

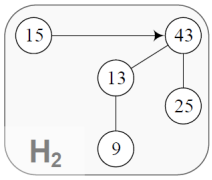
$$h(T_k) = \begin{cases} 0 & k \leq 1 \\ h(T_{k-2}) + 1 & k > 1 \end{cases}$$

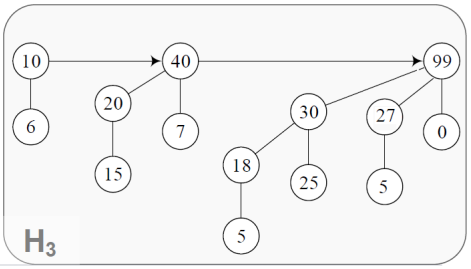
which gives $h(T_k) = h(T_{k-2}) + 1 = h(T_{k-4}) + 2 = \ldots = h(T_{k-2(k/2)}) + k/2 = k/2$. So we have $h(T_k) = k/2$.

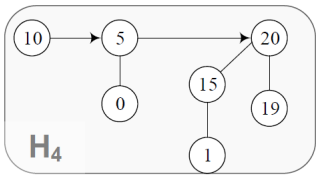
Problem 2 Binomial Heap Operations [4 + 4 + 4 = 12 marks]

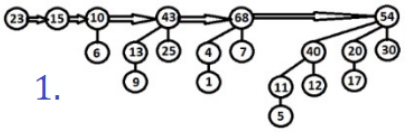
In this problem, we review binomial heap operations for heaps of the figure below. In case of merging trees in the following operations, in case there were three binomial trees of the same order, merge the two 'older' trees (keep the new tree which is the product of previous merge).

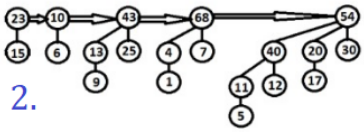


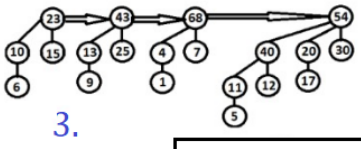


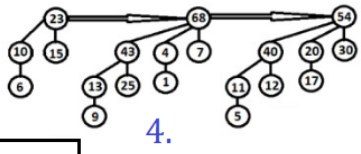


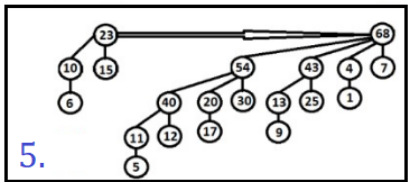










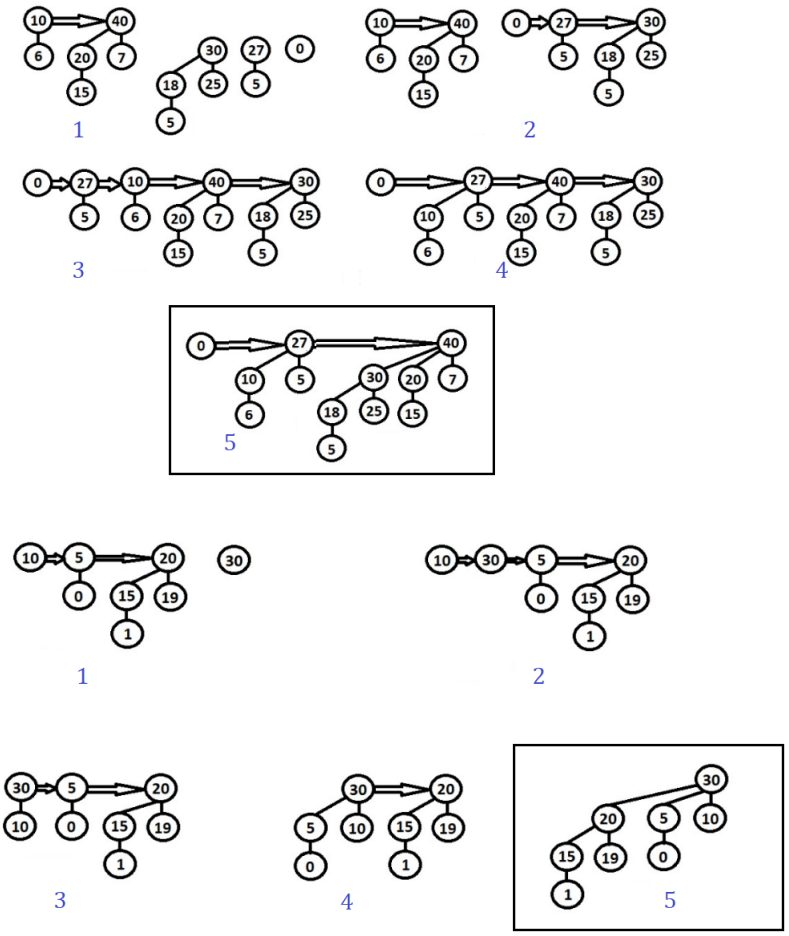


a) Apply the operation merge on heaps H_1 and H_2. Show intermediate steps. **Answer:** See the figure below.

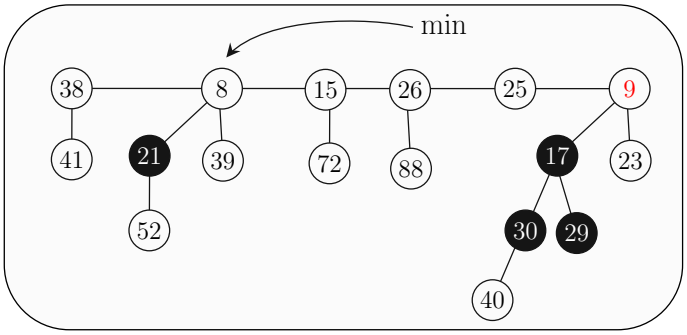
b) Apply operation *extract-max* on heap H_3. Show intermediate steps. **Answer:** See the figure below.

c) Apply operation *insert* on heap H_4. Assume you insert value $x = 30$ to the heap. Show intermediate steps.

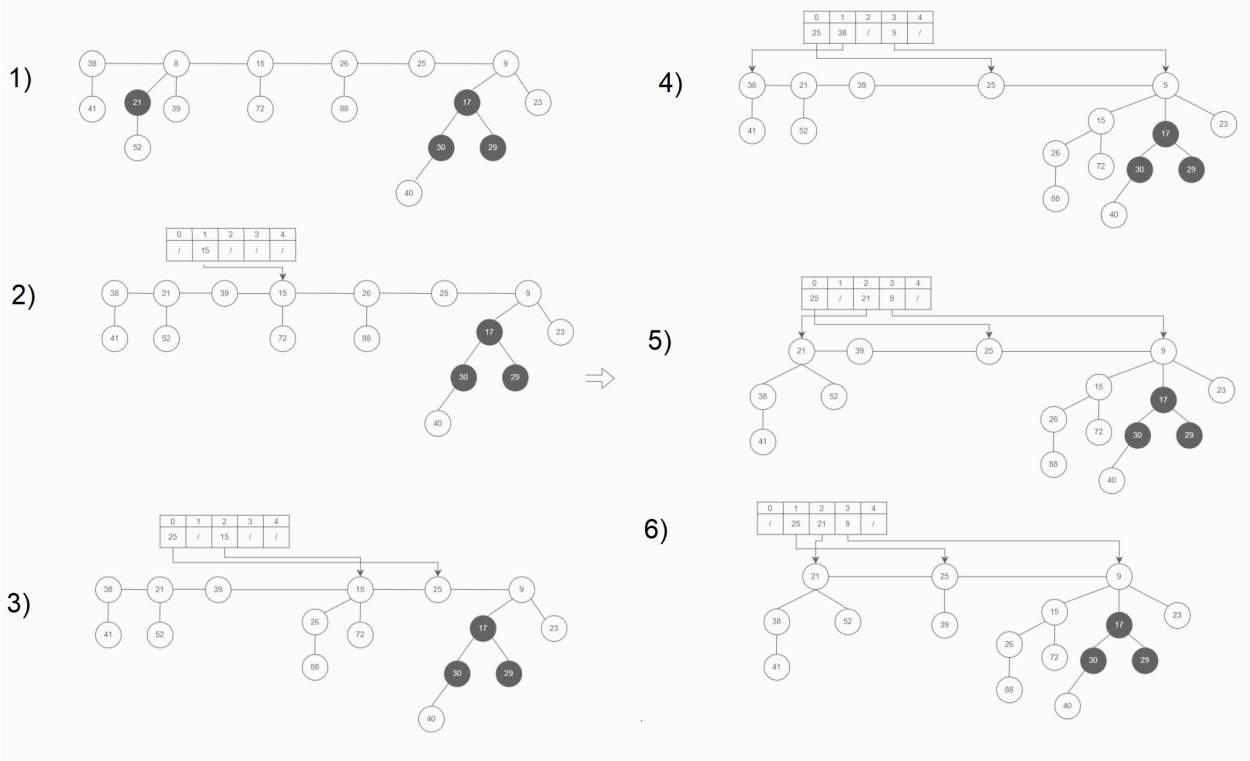
Answer: See the figure.



Problem 3 Fibonacci Heap Operations [4 + 4 = 8 marks]

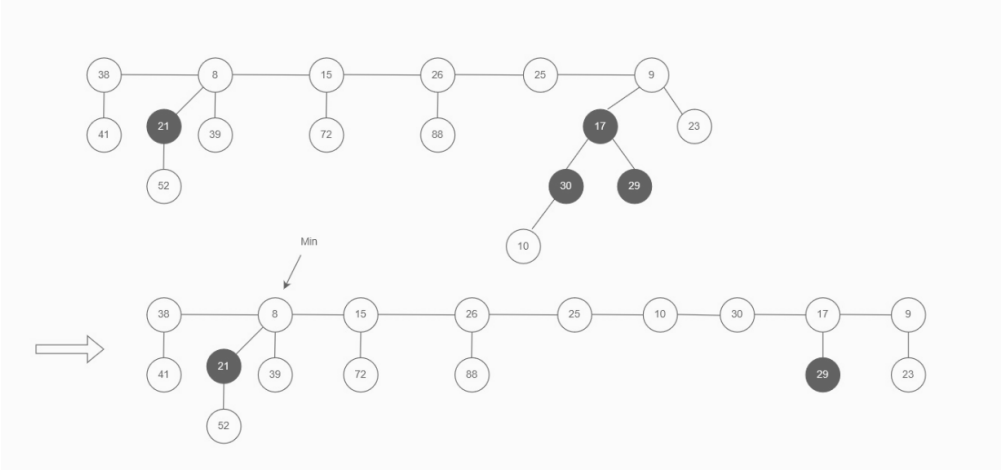


a) In the Fibonacci heap below, apply the operation extractMin. Show your work.
Answer: See the figure below (credit for the the figure goes to one of our great students who submitted it).



b) In the original heap, decrease the key of 40 to 10. Show your work.

Answer: See the figure below:



Problem 4 Fibonacci Heap Analysis [5 + 5 = 10 marks]

Consider a different potential for Fibonacci heaps, defined as $\Phi'(H) = 2t(H) + 3m(H)$. Indicate the amortized time for (a) extractMin and (b) decreaseKey operations. Show your work.

Answer:

- ExtractMin: As before, the actual cost for extractMin is $O(t(H) + \log n)$ (note that the actual cost is independent of the potential). The number of trees is $t(H)$ before extractMin and $O(\log n)$ after the extractMin. As before, the number of mared nodes does not increase. Therefore, we can write down $\Delta(\Phi) \leq 2(O(\log n) - t(H))$, and the amortized cost would be $O(t(H) + \log n) + c(2\log n - 2t(H))$, which is $O(\log n)$, assuming c is large enough.

- DecreaseKey: As before, if the heap property is not violated (case 0), the actual cost is $O(1)$ and the potential does not change, yielding to amortized cost of $O(1)$. Similarly, if the parent of the changed key is unmarked (case 1), actual cost is constant and the potential is increased also by a constant; the amortized cost stays $O(1)$. Finally, suppose the parent of the changed node is marked (case 2), and let p be the number of added trees. Then $t(H)$ is increased by p, which means potential is increased by $2p$. Meanwhile, the number of marked nodes $m(h)$ is decreased by $p-1$, which means the potential is decreased by $3p-3$. The difference in potential is thus $2p-(3p-3) = 3-p$. The actual cost is $O(p)$ and hence the amortized cost will be $O(p) + c(3-p)$, which is $O(1)$, assuming c is large enough.

Problem 5 Union by Weight Analysis [8 marks]

In this problem, we would like to show the amortized time of a union operation when union-by-weight on linked-lists is used is $\Omega(\log n)$. For that, we need to come up with a sequence of $\Theta(n)$ operations for which the amortized cost per operation is $\Omega(\log n)$. We start with *make-set*(x_i) for $i \in \{1, 2, \ldots n\}$ where n is a power of 2. Provide a consequent sequence of $\Theta(n)$ union operations so that the total number of updated pointers for all operations is $\Omega(n \log n)$.

Answer: Here are the sequence of operations and their respective number of pointer-updates:

- Step 1: union$(x_i, x_{i+n/2})$ for $i \in \{1, \ldots, n/2\}$. There will be 1 pointer-update per operation, which sums to $n/2$ total updates in this step.

- Step 2: union$(x_i, x_{i+n/4})$ for $i \in \{1, \ldots, n/4\}$. There will be 2 pointer-update per operation, which sums to $n/4 \times 2 = n/2$ total updates in this step.

- Step 3: union$(x_i, x_{i+n/8})$ for $i \in \{1, \ldots, n/8\}$. There will be 4 pointer-update per operation, which sums to $n/8 \times 4 = n/4$ total updates in this step.

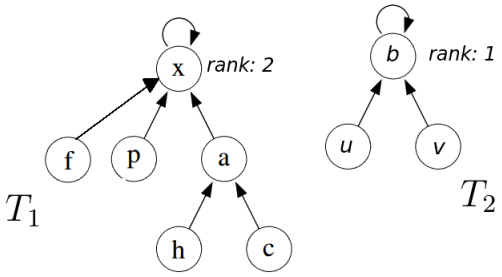
- $\ldots$

- Step k: union$(x_i, x_{i+n/2^k})$ for $i \in \{1, \ldots, n/2^k\}$. There will be 2^{k-1} pointer-update per operation, which sums to $n/2^k \times 2^{k-1} = n/2$ total updates in this step.

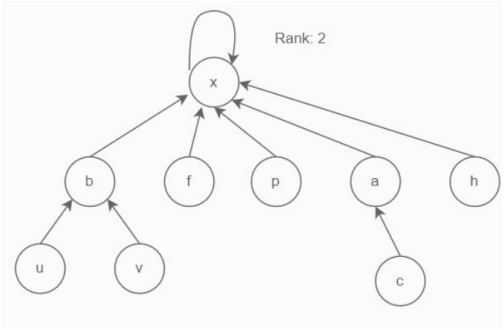
After $k = \lceil \log n \rceil$ steps, there will be most two sets, which are united with union(x_1, x_2), with $n/2$ pointer updates. In summary, there will be roughly $\log n$ steps, each involving update of $n/2$ pointers. The total number of pointer-updates will be $\Omega(n \log n)$.

Problem 6 Union-Find Operations [4 + 4 = 8 marks]

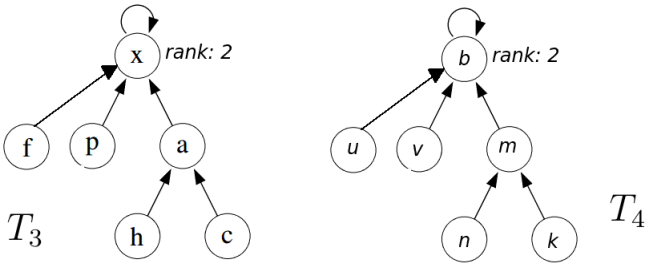
a) Consider a union-find structure based on union-by-rank and path-compression which is formed by T_1 and T_2 in the following figure. Draw the result after the following operations: union(p, v), find(h).



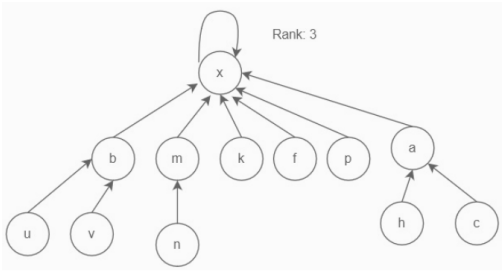
Answer: See the figure below (credit for the figure goes to one of our great students who submitted it). The ordering of the children of the root does not matter.



b) Consider a similar structure formed by T_3 and T_4 in the following figure. Draw the result after the following operations: union(p, u), find(k). For the union operation, as both trees have the same rank, assume x becomes the parent of the united tree.



Answer: See the figure below (credit for the figure goes to one of our great students who submitted it). The ordering of the children of the root does not matter.

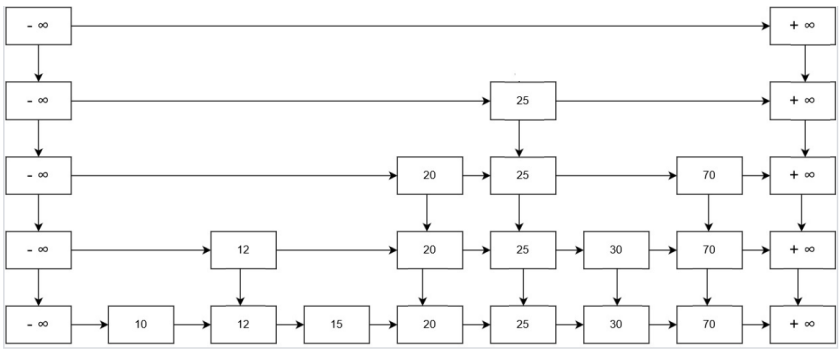


Problem 7 Skip Lists [4 + 6 = 10 marks]

a) Starting with an empty skip list, insert the seven keys $30, 10, 25, 20, 15, 70, 12$. Draw your **final answer** as we saw in the slides. Use the following coin tosses to determine the heights of towers (note, not every toss is necessarily used):

$$H, T, T, H, H, H, T, H, H, T, T, H, H, T, H, T, T, H, T, H, H, T, T, H, H, H, T, H, T, \ldots$$

Answer: See the figure below:



b) Consider a skip list in which we build new towers with probability 3/4.

> When adding an element to the skip list, we flip two coins at the same time, and repeat this until both coins come up tails. The number of times we toss both coins defines the height of the tower (i.e., if both coins come up tails in the first flips, there will be one node in the skip list, if both come coin on the second try, the number of nodes will be 2, etc.).

Using the probability for tower heights described in the above quote, derive the expected height of a tower.

Answer: The height of a tower is 1 if both flips are tails after the first flip of coins; that has a probability of 1/4. The height of a tower is 2 if the first two flips are not both head (chance of 3/4) and in the next trial both flips are tail (chance of 1/4); so the height is 2 with a chance of $3/4 \times 1/4 = 3/16$. Similarly, we have a height h with a chance of $(3/4)^{h-1}$ (not seeing a tail in previous $h-1$ trials) times 1/4 (seeing a tail in h'th trial).

The expected height will be at most

$$\begin{aligned} H &= \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{4} \cdot 2 + (\frac{3}{4})^2 \cdot \frac{1}{4} \cdot 3 + \ldots + (\frac{3}{4})^{h-1} \cdot \frac{1}{4} \cdot h + \ldots \\ &= \frac{1}{4} \cdot (1 + \frac{3}{4} \cdot 2 + \frac{9}{16} \cdot 3 + \frac{27}{64} \cdot 4 + \frac{81}{256} \cdot 5 + \ldots) \end{aligned}$$

Multiplying both sides by 3/4, we get:

$$\frac{3}{4} \cdot H = \frac{1}{4} \cdot (\frac{3}{4} \cdot 1 + \frac{9}{16} \cdot 2 + \frac{27}{64} \cdot 3 + \frac{81}{128} \cdot 4 + \ldots)$$

Hence, we have

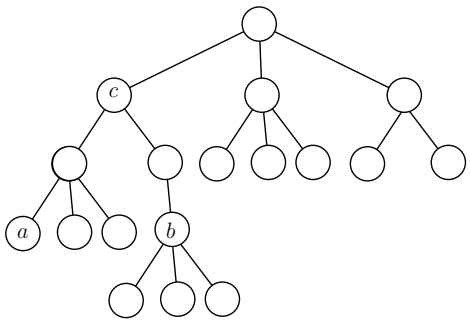
$$\begin{aligned} H - \frac{3}{4} \cdot H &= \frac{1}{4} \cdot (1 + \frac{3}{4} + \frac{9}{16} + \frac{27}{64} + \frac{81}{256} + \ldots) \\ &= \frac{1}{4} \cdot \frac{1}{1 - 3/4} \\ &= \frac{1}{4} \cdot 4 = 1 \end{aligned}$$

So, we have $H - \frac{3}{4} \cdot H = 1$ which implies $H = 4$.

Problem 8 Bonus [7 marks]

Given a rooted tree T, the *deepest common forefather (DCF)* of two nodes u and v is the deepest node x in the tree such that u and v belong to the subtree rooted at x. For example, DCF(a, b) in the figure below is node c.

Assume T is a tree of size n (for some large n) and suppose we are given n pairs of nodes from T. Describe an algorithm that reports the DCF of all pairs in time $o(n \log n)$. Briefly justify your answer.



Answer: The right term for "deepest common forefather (DCF)" is the "lowest common ancestor (LCA)". We changed the name to make it more difficult to search for online answers. Here is a code to find all LCAs (find more details at here):

```
LCA(u)
 1  MAKE-SET(u)
 2  FIND-SET(u).ancestor = u
 3  for each child v of u in T
 4      LCA(v)
 5      UNION(u, v)
 6      FIND-SET(u).ancestor = u
 7  u.color = BLACK
 8  for each node v such that {u, v} ∈ P
 9      if v.color == BLACK
10          print "The least common ancestor of"
                u "and" v "is" FIND-SET(v).ancestor
```

The idea in the code is to traverse the tree and color nodes black as we traverse them. All visited node so far will be maintained in a disjoint-set structure. When we are at node u, the ancestors of u will be stored in different disjoint sets, while its visited (black) descendants are maintained in a set that is represented by u (line 6 ensures that u is the representative). Therefore, as soon as two nodes in a query become members of the same set (as checked in line 8), the representative of that set is reported as their LCA in line 9.