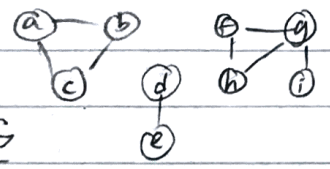
CSC263 Friday Tutorial — Feb 16, 2018

Recall: **Disjoint Sets Data Structure**

- maintains collection $S = \{S_1, S_2, \ldots, S_k\}$ of disjoint sets
- each set identified by representative element
- Operations:
  MAKESET(x), UNION(x,y), FINDSET(x)

- Seems unintuitive data structure; today we will be learning about its uses. Generally used by other algos.

Ex
- Determine the connected components of a graph.



G

$V = \{a, b, c, d, e, f, g, h, i\}$
$E = \{(a,b)\ (b,c)\ (c,d)\ (d,e)\ (f,h)\ (f,g)\ (g,h)\ (g,i)\}$

```
CONNECTED-COMPONENTS(G):
    for each v ∈ G.V: MAKESET(v)
    for each (u,v) ∈ G.E:
        x = FINDSET(u), y = FINDSET(v)
        if x ≠ y: UNION(x,y)
SAME-COMPONENTS(u,v):
    if FINDSET(u) == FINDSET(v) return TRUE;
    else return FALSE
```

Create Disjoint Set Data structure first

Query Data Structure

Recall:
- Linked list implementation of disjoint sets
  - FINDSET $\in O(1)$ - UNION $\in O(n)$
- weighted union heuristic: smaller attached to larger

- Disjoint sets are usually made then extensively explored, ∴ we talk about proving time of sequences of operations...

Prove: A sequence of $n-1$ UNIONs and m FINDSETs is $O(m+n\log n)$.

What is the upperbound on the number of times an element's set representative must be updated?

Consider an element x.
Any time that x's representative is updated, x must have smaller set.
$\therefore$ first time rep updated, new set size ≥ 2
$\therefore$ second time updated, new set size ≥ 4
$\therefore$ i^{th} time updated, new set size $\geq 2^i$
$n \geq 2^i \therefore i \leq \lceil \log_2 n \rceil$

Then representative of any element x updated at most $\log_2 n$.
Then total number of representative updates $\leq n\log_2 n$.
Then UNION $\in O(n\log n)$ (max of sequence of n unions)
Note that m FINDSETs $\in O(m)$.
Then all together $n-1$ unions, m findsets $\in O(m+n\log n)$.

Prove: A sequence of $n-1$ UNIONs and m FINDSETs is $O(m+n\log n)$.

What is the upperbound on the number of times an element's set representative must be updated?

Consider an element x.
Any time that x's representative is updated, x must have smaller set.
$\therefore$ first time rep updated, new set size ≥ 2
$\therefore$ second time updated, new set size ≥ 4
$\therefore$ i^{th} time updated, new set size $\geq 2^i$
$n \geq 2^i \therefore i \in \lceil \log_2 n \rceil$

Then representative of any element x updated at most $\log_2 n$.
Then total number of representative updates $\leq n\log_2 n$.
Then UNION $\in O(n\log n)$ (max of sequence of n unions)
Note that m FINDSETs $\in O(m)$.
Then all together $n-1$ unions, m findsets $\in O(m+n\log n)$.