## §4 Fast&Slowly Growing Functions

# §4 Disjoint-Set Data Structure

**MakeSet**(x) **FindSet**(x) **Union**(x,y) return "handle" **Goal:** amortized  $O^*(1)$ 

**Example Application:** graph G=(V,E) connected components with 'growing' *E*:

sameComponent(u,v): return FindSet(u)=FindSet(v)
addEdge(u,v): Union(u,v)

Naive implementation as forest of depth 1: n calls, m:=n/2 of which are MakeSet  $\rightarrow$  runtime quadratic in n*Weighted union heuristic*: attach smaller to larger tree

**Theorem:** This yields total time  $O(n+m \cdot \log m)$  for any sequence of *n* calls, *m* of which are MakeSet

### §4 Analysis of Union-by-Weight



**MakeSet**(x) **FindSet**(x) **Union**(x,y)



**Observation:** An element's link is updated only when its set is combined with one of more or equal weight.

So to any of the  $\leq m$  possible elements x, k updates occur only <u>after</u> having made  $m \geq 2^k$  calls to <u>MakeSet</u> MakeSet, FindSet: O(1), Union ?

Weighted union heuristic: attach smaller to larger tree

**Theorem:** This yields total time  $O(n+m \cdot \log m)$  for any sequence of *n* calls, *m* of which are MakeSet



Naive implementation as forest of <u>unbounded</u> depth:

MakeSet, Union: O(1), FindSet ? Path compression Lazy Union-by-rank: attach shallower to deeper tree

**Theorem:** This algorithm makes m MakeSet and n-mFindSet and Union calls run in total time  $O(n \cdot \alpha(m))$ .

#### Amortized Analysis of Union-Find



**FindSet**(x) **Union**(x,y)

CS500 M. Ziegler

**Claim a)** Ranks increase strictly along each path.

**b)** Any node of rank r is ancester to  $\geq 2^r$  nodes.

**c)** No more than  $m/2^r$  nodes can have rank  $\geq r$ 

 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1
 1

**Theorem:** This algorithm makes m MakeSet and n-mFindSet and Union calls run in total time  $O(n \cdot \log^* m)$ .

## **Review of Chapters 1 to 4**



- Virtues of Theoretical Computer Science:
  - full and unambiguous problem specification
  - formal semantics of primitive operations
  - algorithm design (as opposed to 'programming')
  - and analysis (correctness, asymptotic cost)
  - optimality proof
- Stable Matchings and Gale-Shapley Algorithm
- AVL-Trees, Binomial Heaps
- Amortized Analysis, Potential Method
- Fibonacci Heaps, Relaxed Binomial Trees
- Fast/Slowly Growing Functions
- Union-Find Algorithm&Analysis