



Algorithms: Design
and Analysis, Part II

Advanced Union-Find

Union by Rank -
Analysis

Properties of Ranks

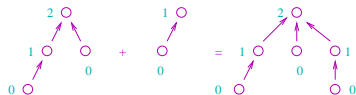
Recall: Lazy Unions.

Invariant (for now): $\text{rank}[x] = \max \# \text{ of hops from a leaf to } x$.

[Note $\max_x \text{rank}[x] \approx \text{worst-case running time of FIND.}$]

Union by Rank: Make old root with smaller rank child of the root with the larger rank.

[Choose new root arbitrarily in case of a tie, and add 1 to its rank.]



Immediate from Invariant/Rank Maintenance:

- (1) For all objects x , $\text{rank}[x]$ only goes up over time
- (2) Only ranks of roots can go up
[once x a non-root, $\text{rank}[x]$ frozen for evermore]
- (3) Ranks strictly increase along a path to the root

Rank Lemma

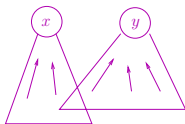
Rank Lemma: Consider an arbitrary sequence of UNION (+FIND) operations. For every $r \in \{0, 1, 2, \dots\}$, there are at most $n/2^r$ objects with rank r .

Corollary: Max rank always $\leq \log_2 n$

Corollary: Worst-case running time of FIND, UNION is $O(\log n)$.
[With Union by Rank.]

Proof of Rank Lemma

Claim 1: If x, y have the same rank r , then their subtrees (objects from which can reach x, y) are disjoint.



Claim 2: The subtree of a rank- r object has size $\geq 2^r$.
[Note Claim 1 + Claim 2 imply the Rank Lemma.]

Proof of Claim 1: Will show contrapositive. Suppose subtrees of x, y have object z in common $\Rightarrow \exists$ path $z \rightarrow x, z \rightarrow y$
 \Rightarrow One of x, y is an ancestor of the other
 \Rightarrow The ancestor has strictly larger rank. [By property (3)]

QED (Claim 1)

Proof of Claim 2

Rank $r \Rightarrow$ Subtree size $\geq 2^r$

Base case: Initially all ranks = 0, all subtree sizes = 1

Inductive step: Nothing to prove unless the rank of some object changes (subtree sizes only go up).

Interesting case: UNION(x, y), with $s_1 = \text{FIND}(x)$, $s_2 = \text{FIND}(y)$, and $\text{rank}[s_1] = \text{rank}[s_2] = r \Rightarrow s_2$'s new rank = $r + 1$
 $\Rightarrow s_2$'s new subtree size = s_2 's old subtree size + s_1 's old subtree size (each at least 2^r by the inductive hypothesis) $\geq 2^{r+1}$. QED!

