Union Find

- Union Find
- Quick Find
- Quick Union
- Weighted Quick Union
- Path Compression
- Dynamic Connectivity

Philip Bille

Union Find

- Union Find
- Quick Find
- Quick Union
- Weighted Quick Union
- Path Compression
- Dynamic Connectivity

Philip Bille

Union Find

- Union find. Maintain a dynamic family of sets supporting the following operations:
  - INIT(n): construct sets {0}, {1},…, {n-1}
  - UNION(i,j): forms the union of the two sets that contain i and j. If i and j are in the same set nothing happens.
  - FIND(i): return a representative for the set that contains i.

INIT(9)

{0} {1} {2} {3} {4} {5} {6} {7} {8}

{1, 0, 6} {8, 3, 2, 7} {4, 5} $\xrightarrow{\text{UNION}(5,0)}$ {1, 0, 6, 4, 5} {8, 3, 2, 7}

Union Find

- Applications.
    - Dynamic connectivity.
    - Minimum spanning tree.
    - Unification in logic and compilers.
    - Nearest common ancestors in trees.
    - Hoshen-Kopelman algorithm in physics.
    - Games (Hex and Go)
    - Illustration of clever techniques in data structure design.

Union Find

- Union Find
- Quick Find
- Quick Union
- Weighted Quick Union
- Path Compression
- Dynamic Connectivity

Philip Bille

Quick Find

- Quick find. Maintain array id[0..n-1] such that id[i] = representative for i.
  - INIT(n): set elements to be their own representative.
  - UNION(i,j): if FIND(i) ≠ FIND(j), update representative for all elements in one of the sets.
  - FIND(i): return representative.

INIT(9)

{0} {1} {2} {3} {4} {5} {6} {7} {8}

	0	1	2	3	4	5	6	7	8
id[]	0	1	2	3	4	5	6	7	8

UNION(5,0)

{1, 0, 6} {8, 3, 2, 7} {4, 5} → {1, 0, 6, 4, 5} {8, 3, 2, 7}

	0	1	2	3	4	5	6	7	8
id[]	1	1	3	3	5	5	1	3	3

0	1	2	3	4	5	6	7	8
1	1	3	3	1	1	1	3	3

Quick Find

```
Init(n):
    for k = 0 to n-1
        id[k] = k
```

```
Find(i):
    return id[i]
```

```
Union(i,j):
    iID = Find(i)
    jID = Find(j)
    if (iID ≠ jID)
        for k = 0 to n-1
            if (id[k] == iID)
                id[k] = jID
```

{1, 0, 6} {8, 3, 2, 7} {4, 5} Union(5,0) → {1, 0, 6, 4, 5} {8, 3, 2, 7}

	0	1	2	3	4	5	6	7	8
id[]	1	1	3	3	5	5	1	3	3

	0	1	2	3	4	5	6	7	8
	1	1	3	3	1	1	1	3	3

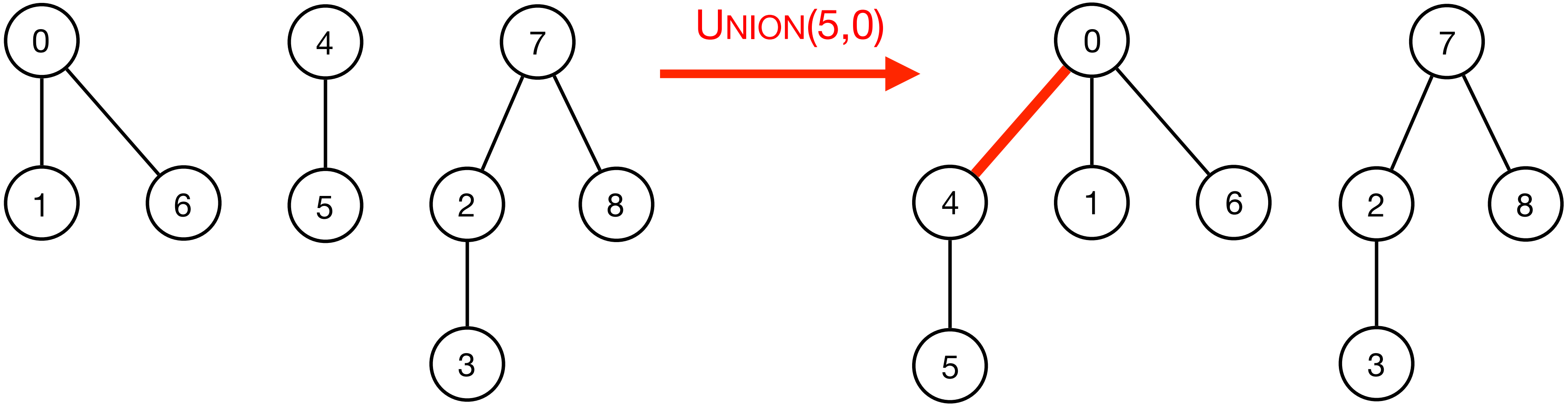
- Time.
    - O(n) time for Init, O(n) time for Union, and O(1) tid for Find.

Union Find

- Union Find
- Quick Find
- Quick Union
- Weighted Quick Union
- Path Compression
- Dynamic Connectivity

Quick Union

- Quick union. Maintain each sets as a rooted tree.
- Store trees as array p[0..n-1] such that p[i] is the parent of i and p[root] = root. Representative is the root of tree.
  - INIT(n): create n trees with one element each.
  - UNION(i,j): if FIND(i) ≠ FIND(j), make the root of one tree the child of the root of the other tree.
  - FIND(i): follow path to root and return root.



	0	1	2	3	4	5	6	7	8
p[]	0	0	7	2	4	4	0	7	7

0	1	2	3	4	5	6	7	8
0	0	7	2	0	4	0	7	7

Quick Union

- INIT(n): create n trees with one element each.
- UNION(i,j): if FIND(i) ≠ FIND(j), make the root of one tree the child of the root of the other tree.
- FIND(i): follow path to root and return root.

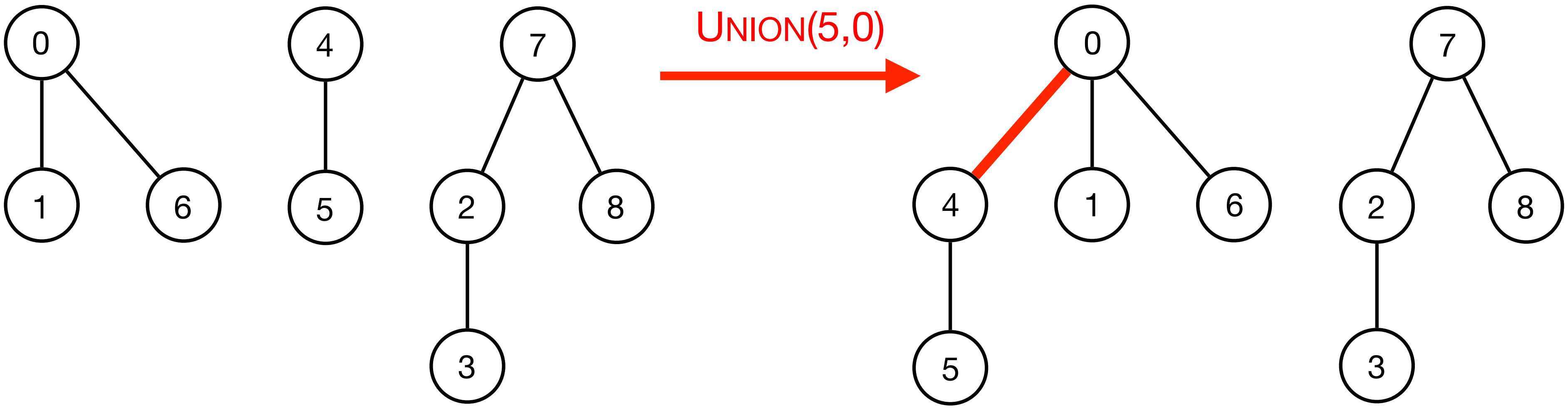
- Exercise. Show data structure after each operation in the following sequence.
  - INIT(7), UNION(0,1), UNION(2,3), UNION(5,1), UNION(5,0), UNION(0,3), UNION(5,2), UNION(4,3), UNION(4,6).

Quick Union

```
INIT(n):
    for k = 0 to n-1
        p[k] = k
```

```
FIND(i):
    while (i != p[i])
        i = p[i]
    return i
```

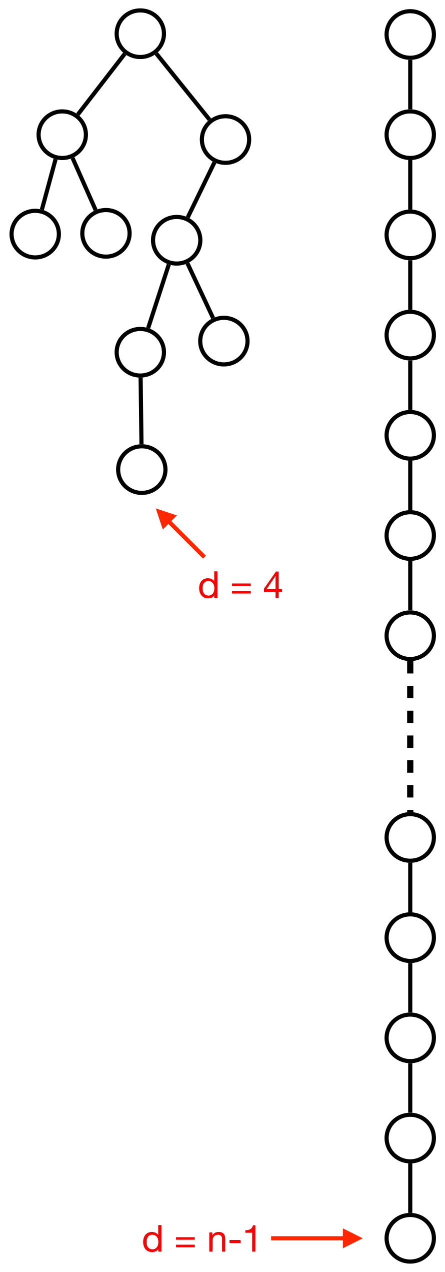
```
UNION(i,j):
    $r_i$ = FIND(i)
    $r_j$ = FIND(j)
    if ($r_i \neq r_j$)
        p[$r_i$] = $r_j$
```



- Time.
    - O(n) time for INIT, O(d) tid for UNION and FIND, where d is the depth of the tree.

Quick Union

- UNION and FIND depend on the depth of the tree.
- Bad news. Depth can be n-1.
- Challenge. Can combine trees to limit the depth?

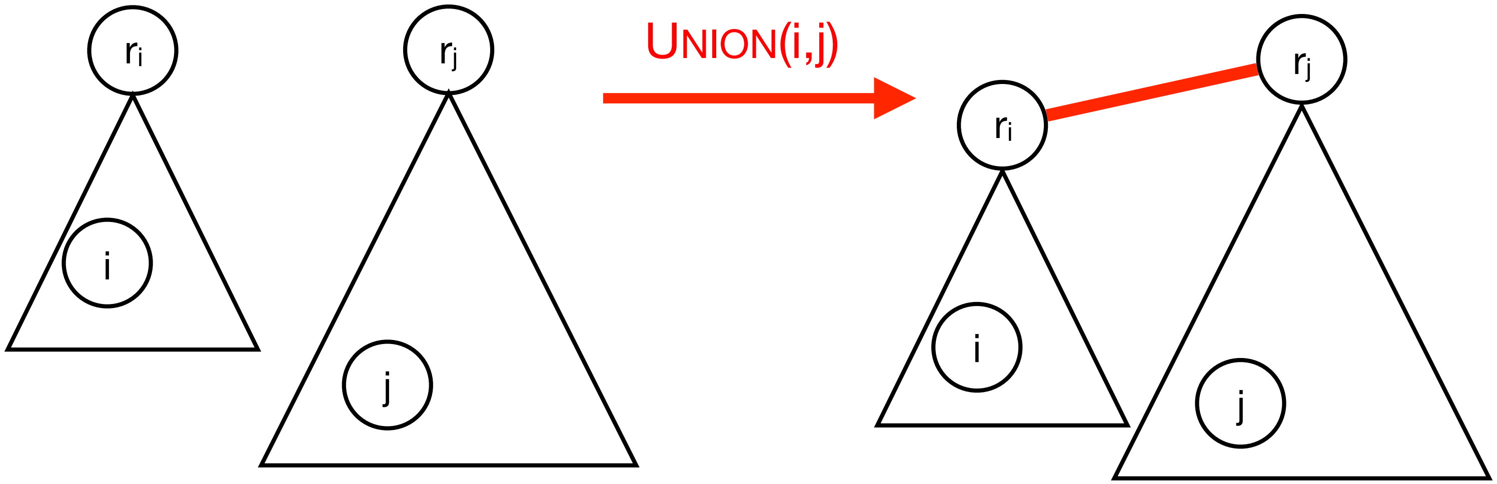


Union Find

- Union Find
- Quick Find
- Quick Union
- Weighted Quick Union
- Path Compression
- Dynamic Connectivity

Weighted Quick Union

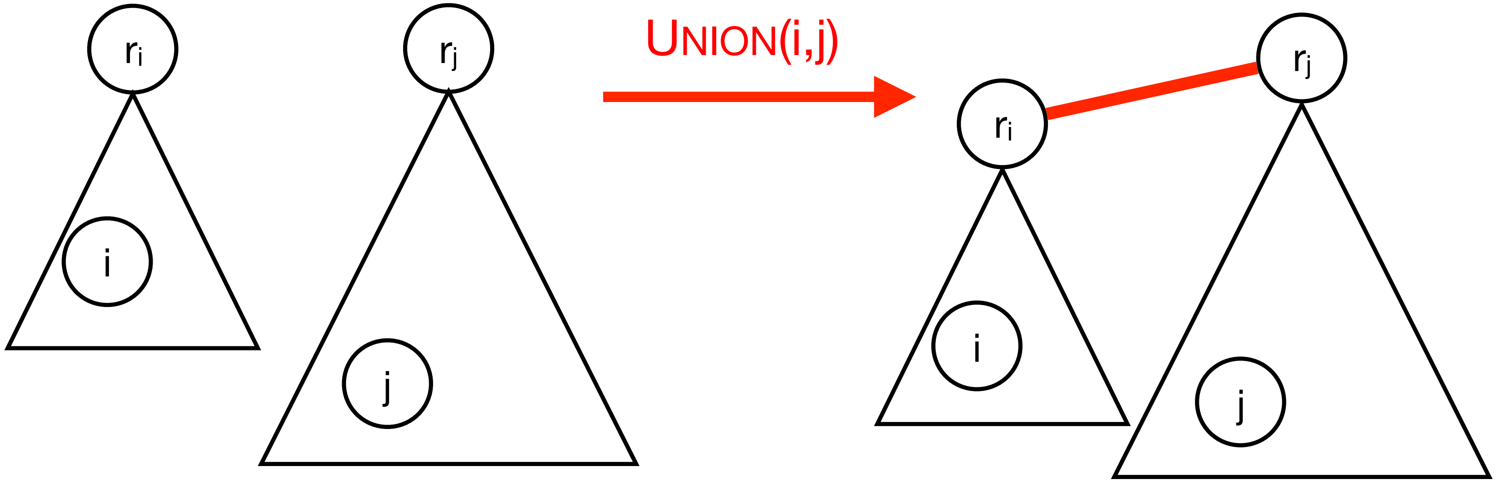
- Weighted quick union. Extension of quick union.
- Maintain extra array sz[0..n-1] such sz[i] = the size of the subtree rooted at i.
  - Init: as before + initialize sz[0..n-1].
  - Find: as before.
  - Union(i,j): if Find(i) ≠ Find(j), make the root of the smaller tree the child of the root of the larger tree.
- Intuition. Union balances the trees.



$sz[r_j] = sz[r_i] + sz[r_j]$

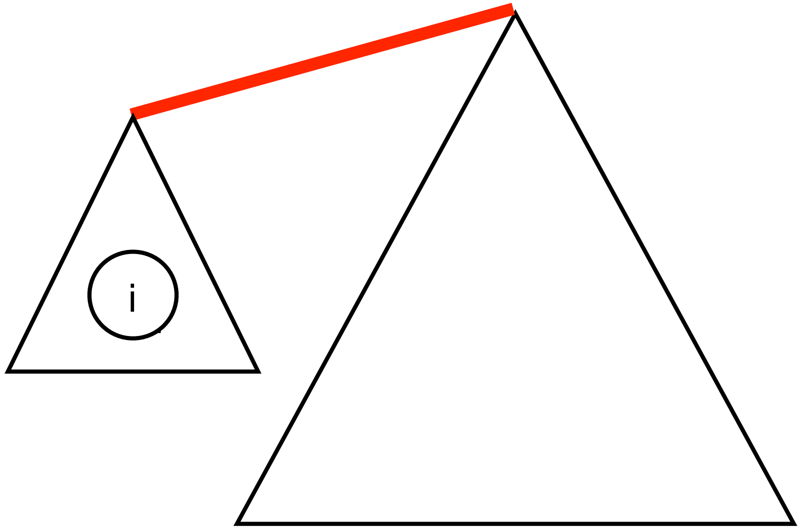
Weighted Quick Union

```
UNION(i,j):
    $r_i$ = FIND(i)
    $r_j$ = FIND(j)
    if ($r_i \neq r_j$)
        if (sz[$r_i$] < sz[$r_j$])
            p[$r_i$] = $r_j$
            sz[$r_j$] = sz[$r_i$] + sz[$r_j$]
        else
            p[$r_j$] = $r_i$
            sz[$r_i$] = sz[$r_i$] + sz[$r_j$]
```



Weighted Quick Union

- Lemma. With weighted quick union the depth of a node is at most $\log_2 n$.
- Proof.
  - Consider node i with depth d_i.
  - Initially $d_i = 0$.
  - d_i increases with 1 when the tree is combined with a larger tree.
  - The combined tree is at least twice the size.
  - We can double the size of trees at most $\log_2 n$ times.
  - $\Longrightarrow d_i \leq \log_2 n$.



Union Find

Data structure	UNION	FIND
quick find	O(n)	O(1)
quick union	O(n)	O(n)
weighted quick union	O(log n)	O(log n)

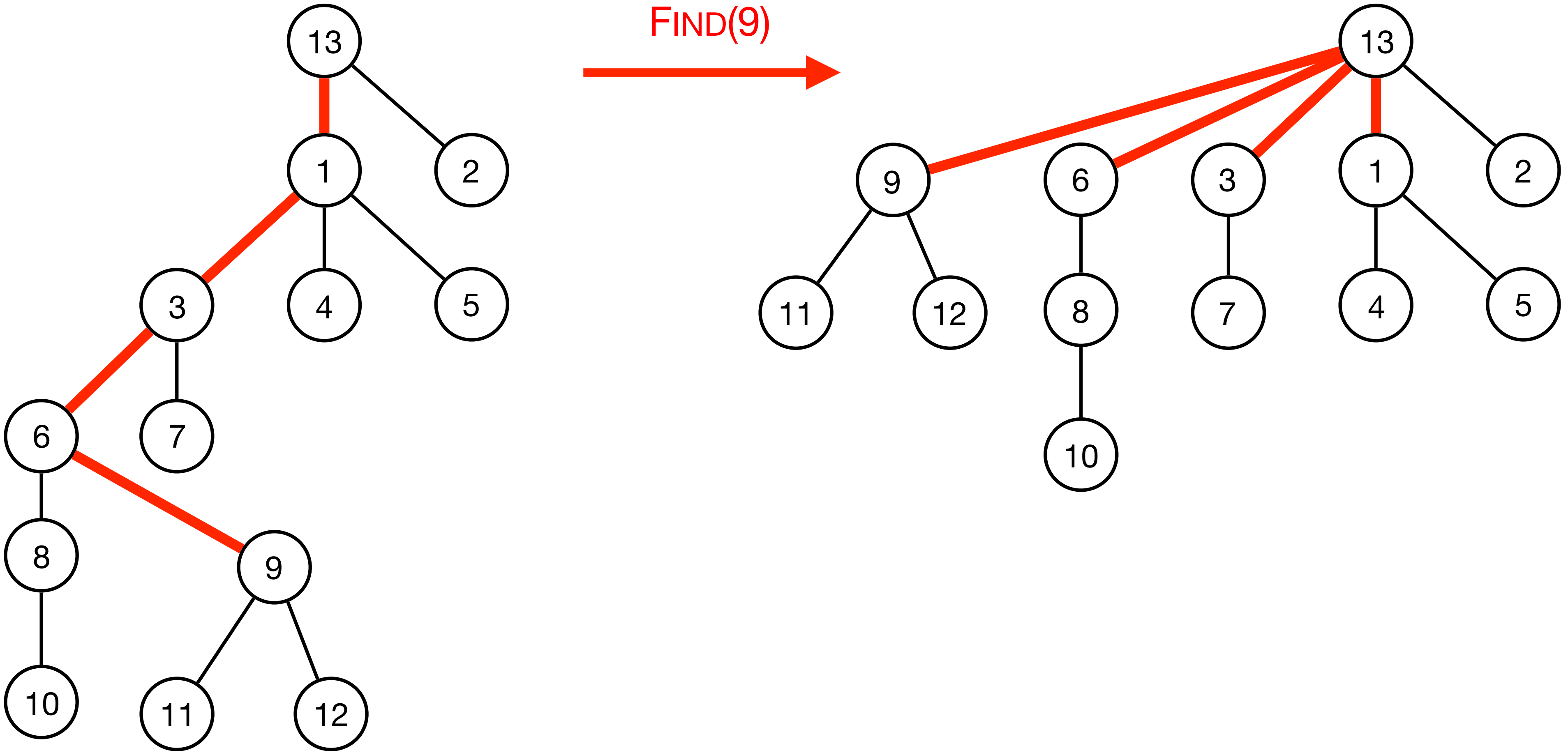
- Challenge. Can we do even better?

Union Find

- Union Find
- Quick Find
- Quick Union
- Weighted Quick Union
- Path Compression
- Dynamic Connectivity

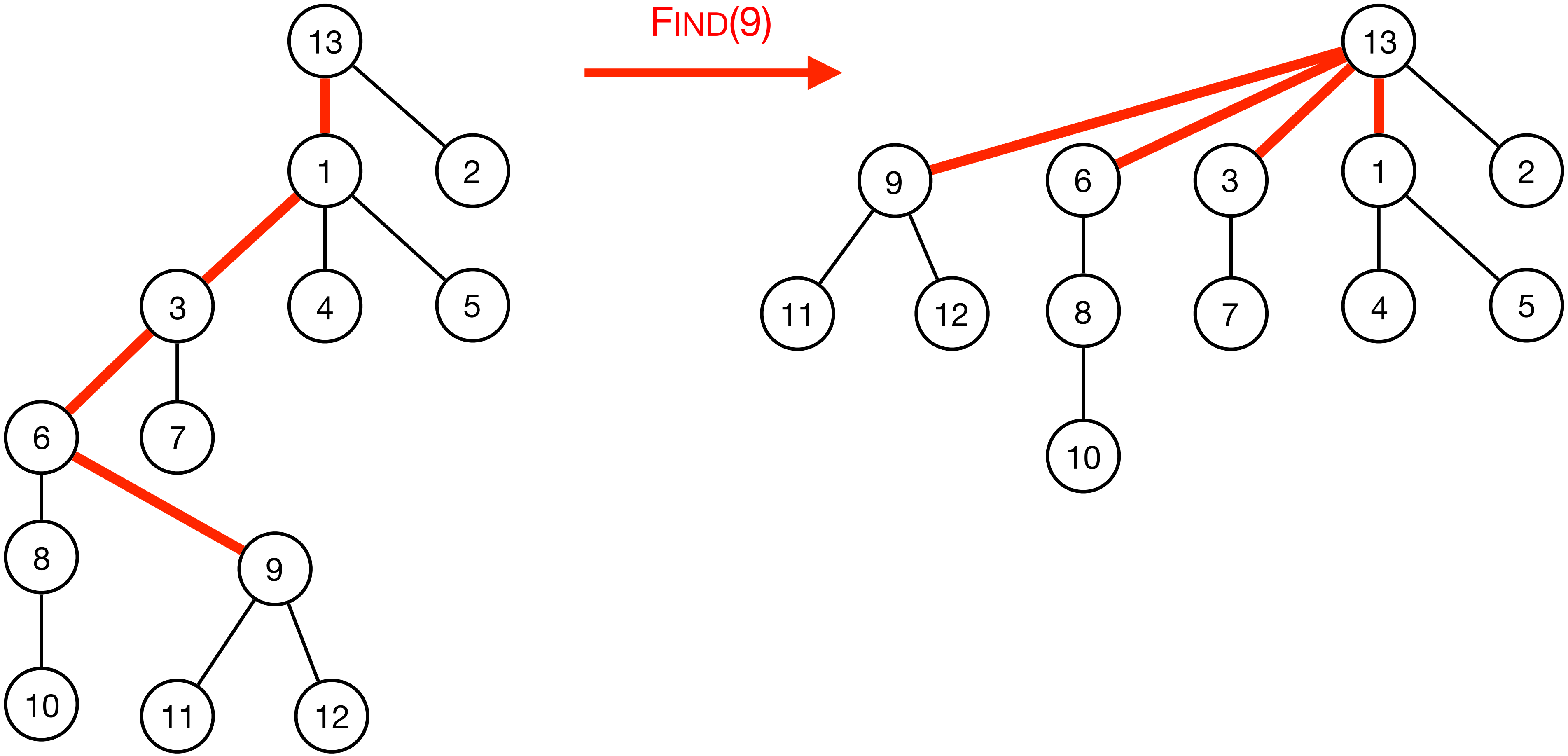
Path Compression

- Path compression. Compress path on Find. Make all nodes on the path children of the root.
- No change in running time for a single Find. Subsequent Find become faster.
- Works with both quick union and weighted quick union.



Path Compression

- Theorem [Tarjan 1975]. With path compression any sequence of m Find og Union operations on n elements take O(n + m α(m,n)) time.
- α(m,n) is the inverse of Ackermanns function. α(m,n) ≤ 5 for any practical input.
- Theorem [Fredman-Saks 1985]. It is not possible to support m Find og Union operations O(n + m) time.

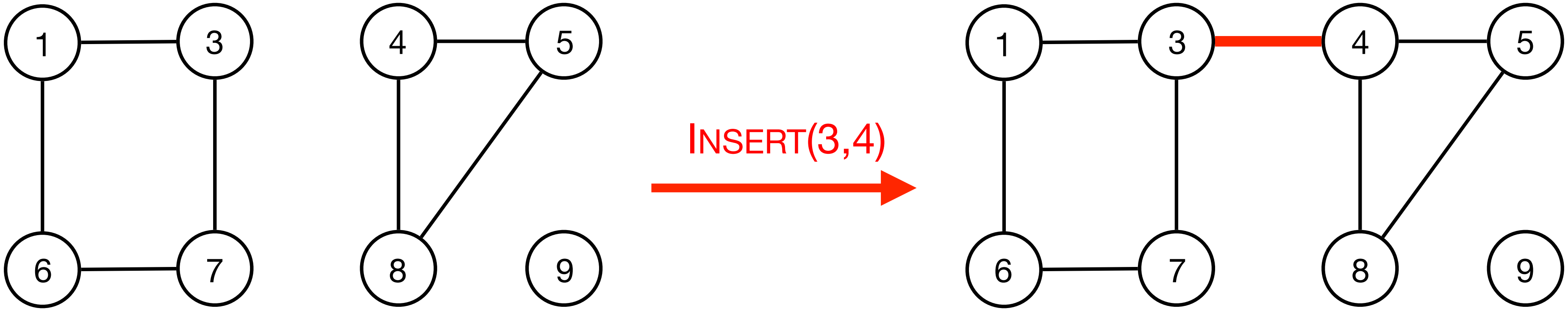


Union Find

- Union Find
- Quick Find
- Quick Union
- Weighted Quick Union
- Path Compression
- Dynamic Connectivity

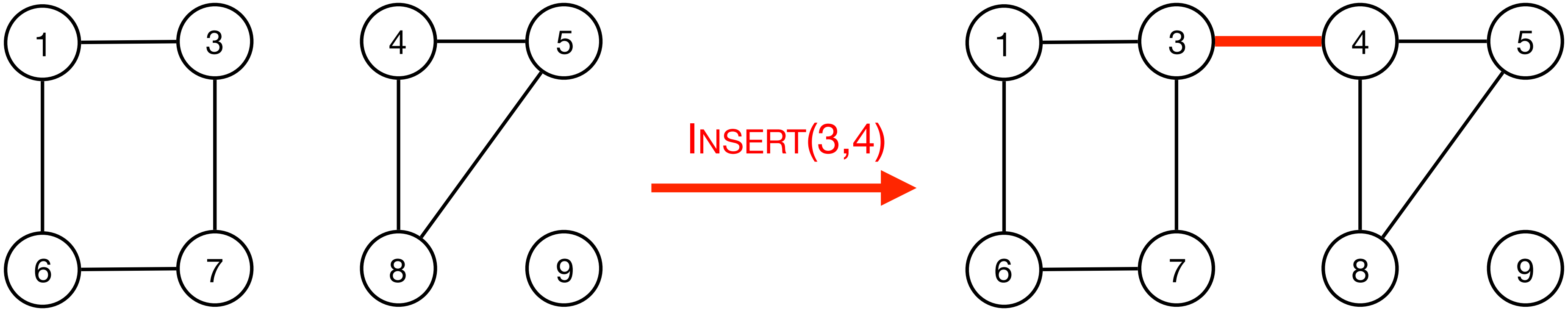
Dynamic Connectivity

- Dynamic connectivity. Maintain a dynamic graph supporting the following operations:
  - INIT(n): create a graph G med n vertices and no edges.
  - CONNECTED(u,v): determine if u og v are connected.
  - INSERT(u, v): add edge (u,v). We assume (u,v) does not already exists.



Dynamic Connectivity

- Implementation with union find.
  - INIT(n): initialize a union find data structure with n elements.
  - CONNECTED(u,v): FIND(u) == FIND(v).
  - INSERT(u, v): UNION(u,v)



- Time
  - O(n) time for INIT, O(log n) time for CONNECTED, and O(log n) time for INSERT

Union Find

- Union Find
- Quick Find
- Quick Union
- Weighted Quick Union
- Path Compression
- Dynamic Connectivity