

Disjoint Set Union (DSU)

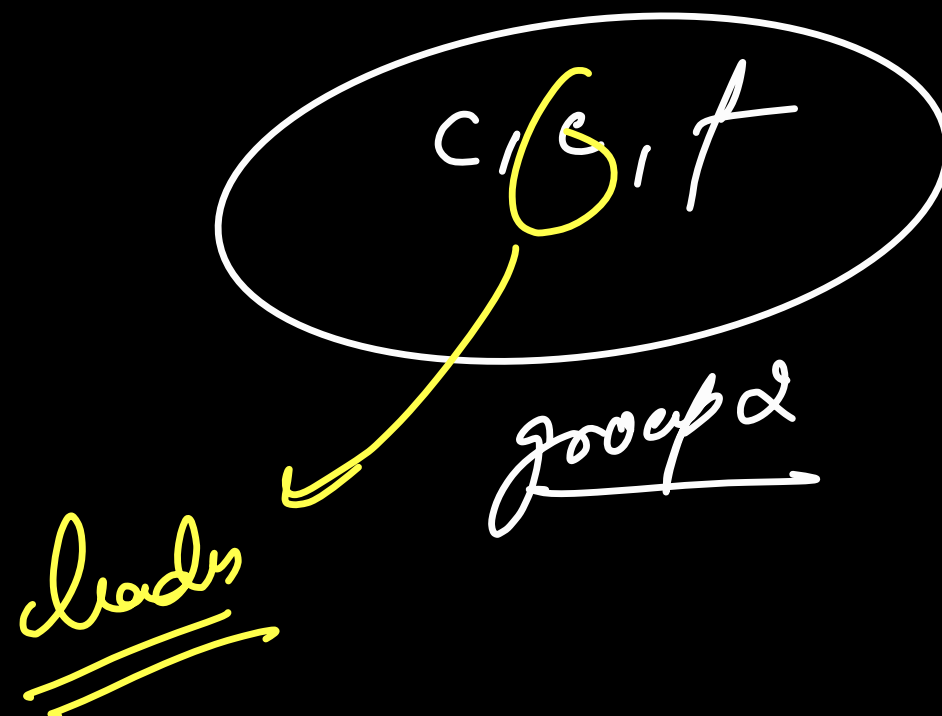
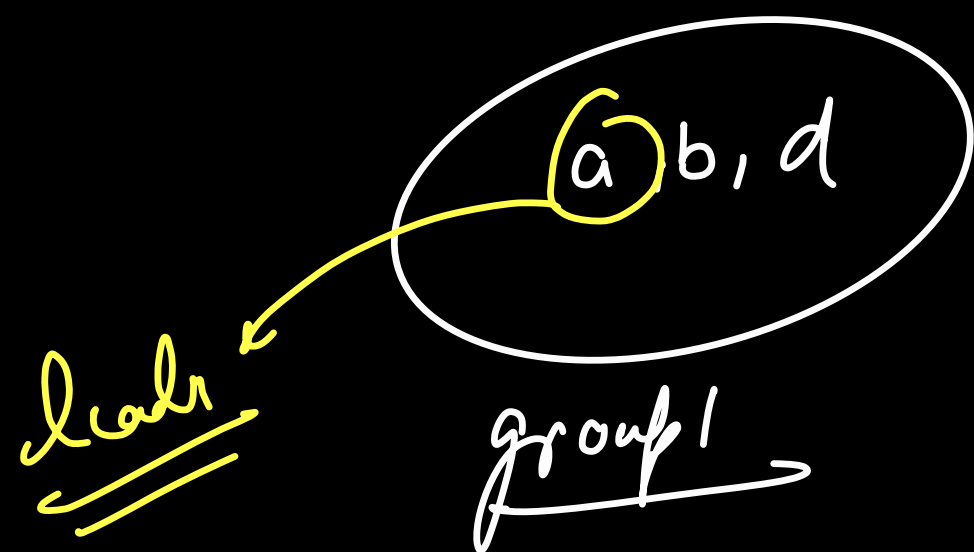
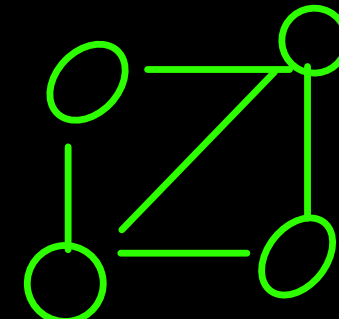
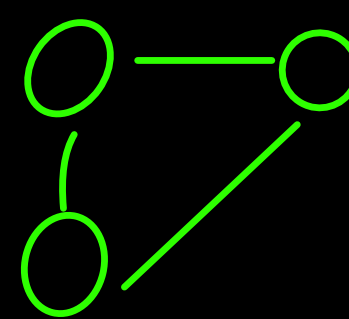
↳ Linear DS → stacks, queues, arrays, ll etc.

↳ Non-linear DS → Hashmaps

↳ Trees → BT, BST, heaps

Clustering / grouping → You will be having some elements in diff groups / & you need to add them / segregate them
clusters. and sometimes we might need to identify the group any element belongs to.

a, b, c, d, ..., f, g, h



Terminologies

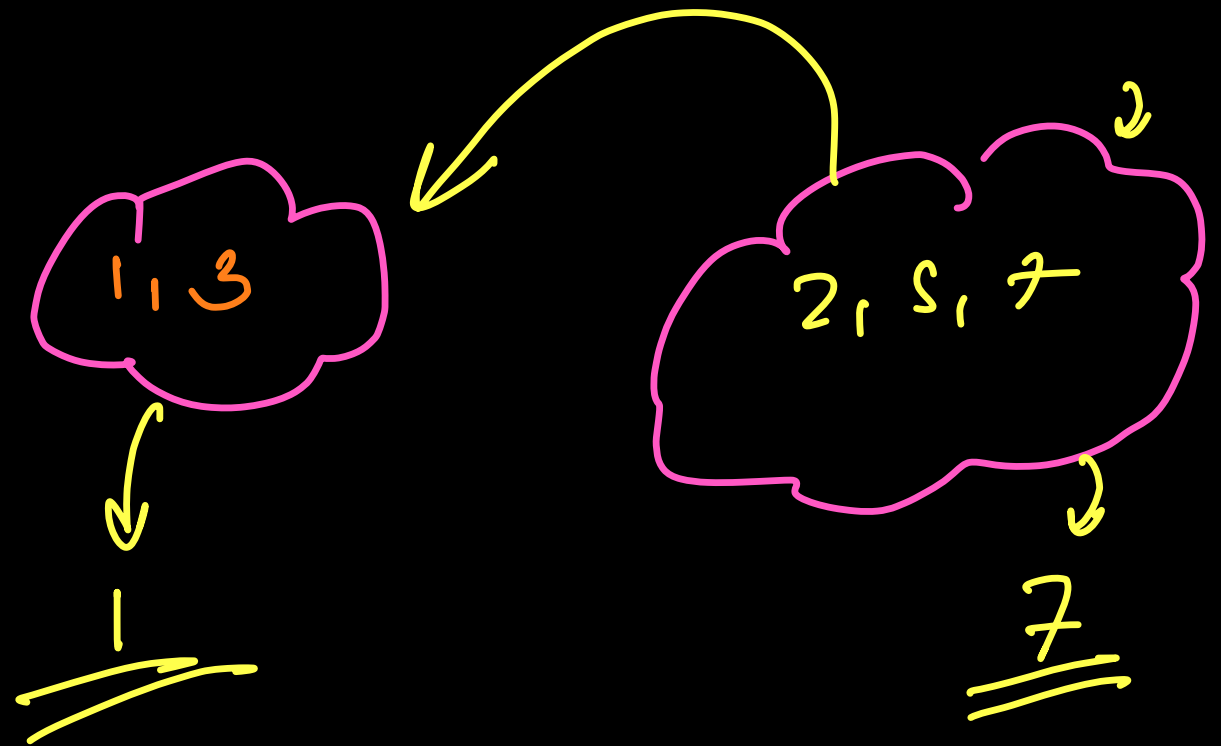
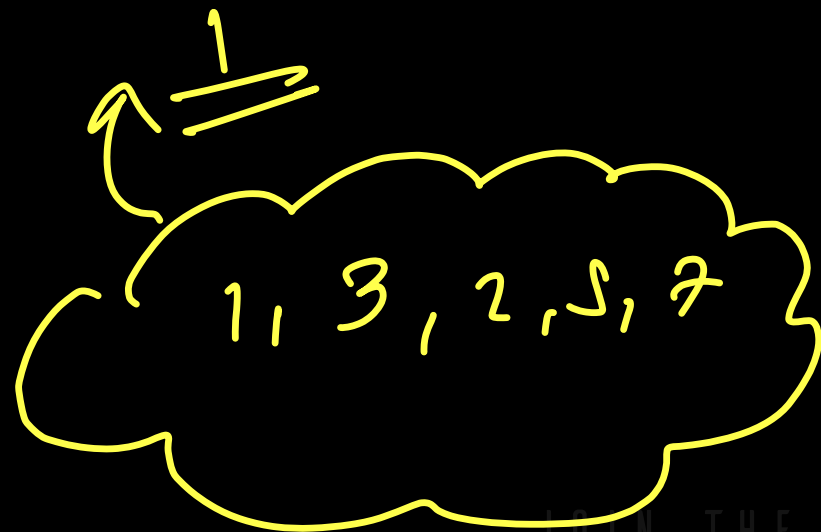
① Leader / parent of the group → to uniquely identify a group we will pick any element from the group & make it the representative / leader / parent of the group.

DSU

Q what funcⁿ dsu need to support ??

① Union (a, b) → adds the group where element b belongs to the group where element a belongs or vice-versa.

Ex → union (1, 5)



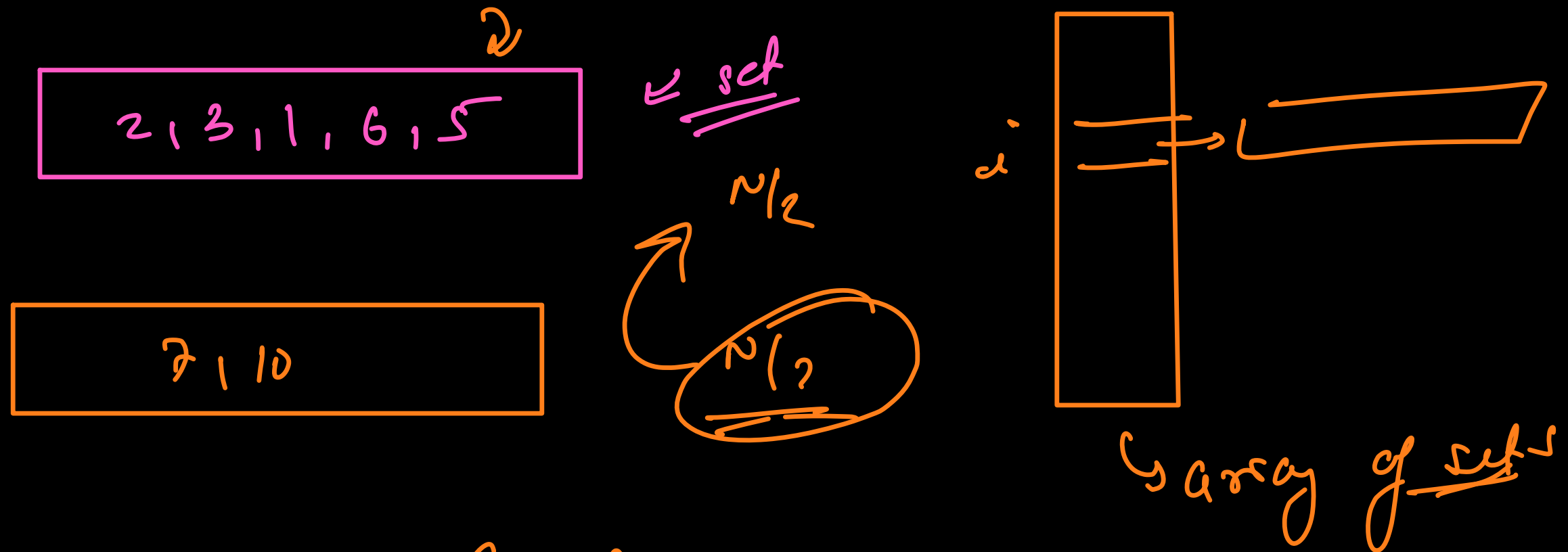
② $\text{find}(x) / \text{get}(x)$: this will be used to find which group x belongs to. we will return the parent of the group that x belongs to.

$\{2, 3, 1, 6\}$ \rightarrow parent = 3

$\text{get}(2) \rightarrow 3$

$\text{get}(3) \rightarrow \underline{\underline{3}}$

Approach \rightarrow Represent every group as a set.



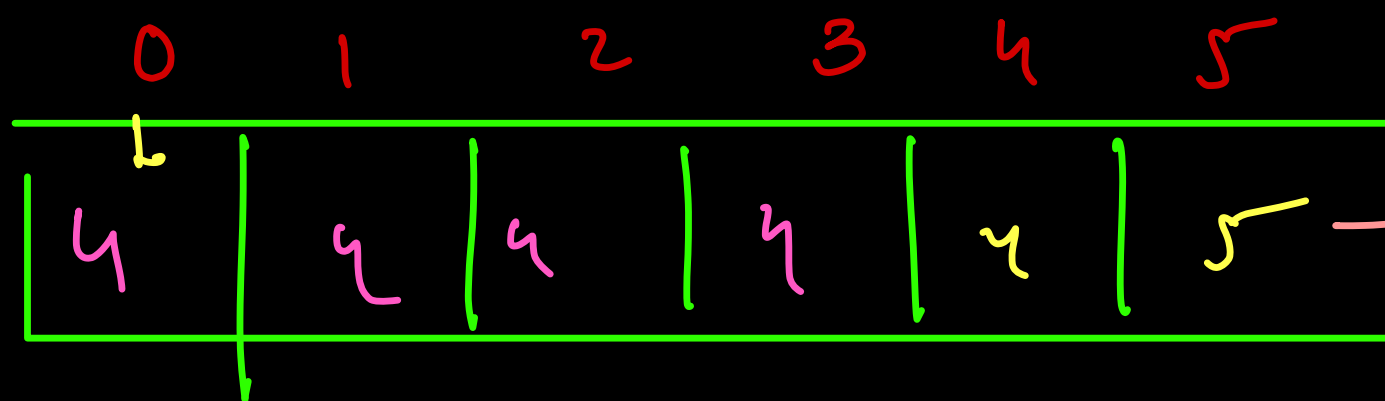
$O(n \log n)$

\hookrightarrow $O(n)$

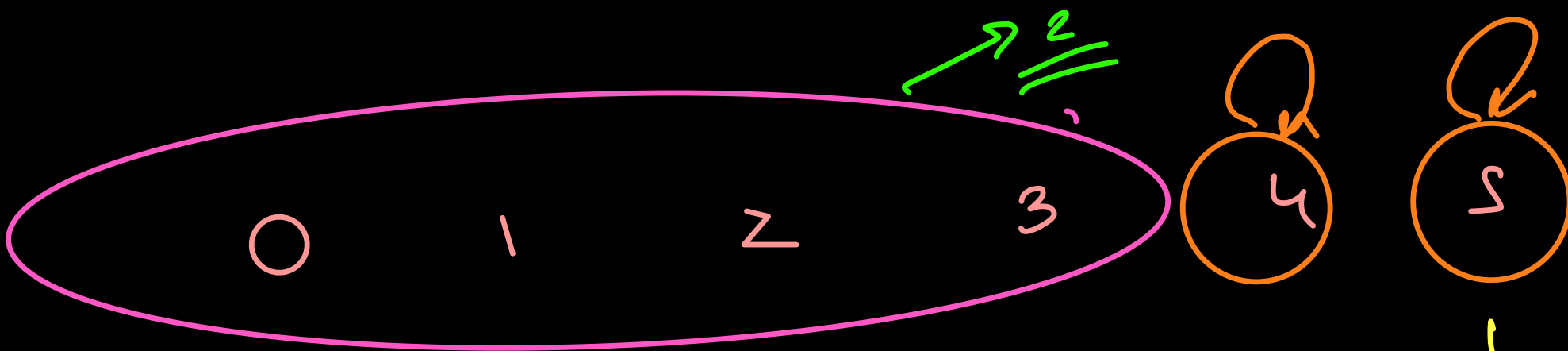
approach 2

Can we use arrays??

indices
represent
actual parent



values \Rightarrow $par[i]$
denotes which
group it belongs



union (0, 1)

union (2, 3)

union (4, 3)

$par[i] \rightarrow p$

union (2, 1)

initially every
element belongs to
diff own groups

```
int find(x) {
    return par[x];
}
```

$\rightarrow \underline{\underline{O(1)}}$

```
void union(a, b) {
    a = find(a)
    b = find(b)
```

$\rightarrow \underline{\underline{O(1)}}$

```
    for (i=0; i<n; i++)
        if (par[i] == b)
            par[i] = a
    }
```

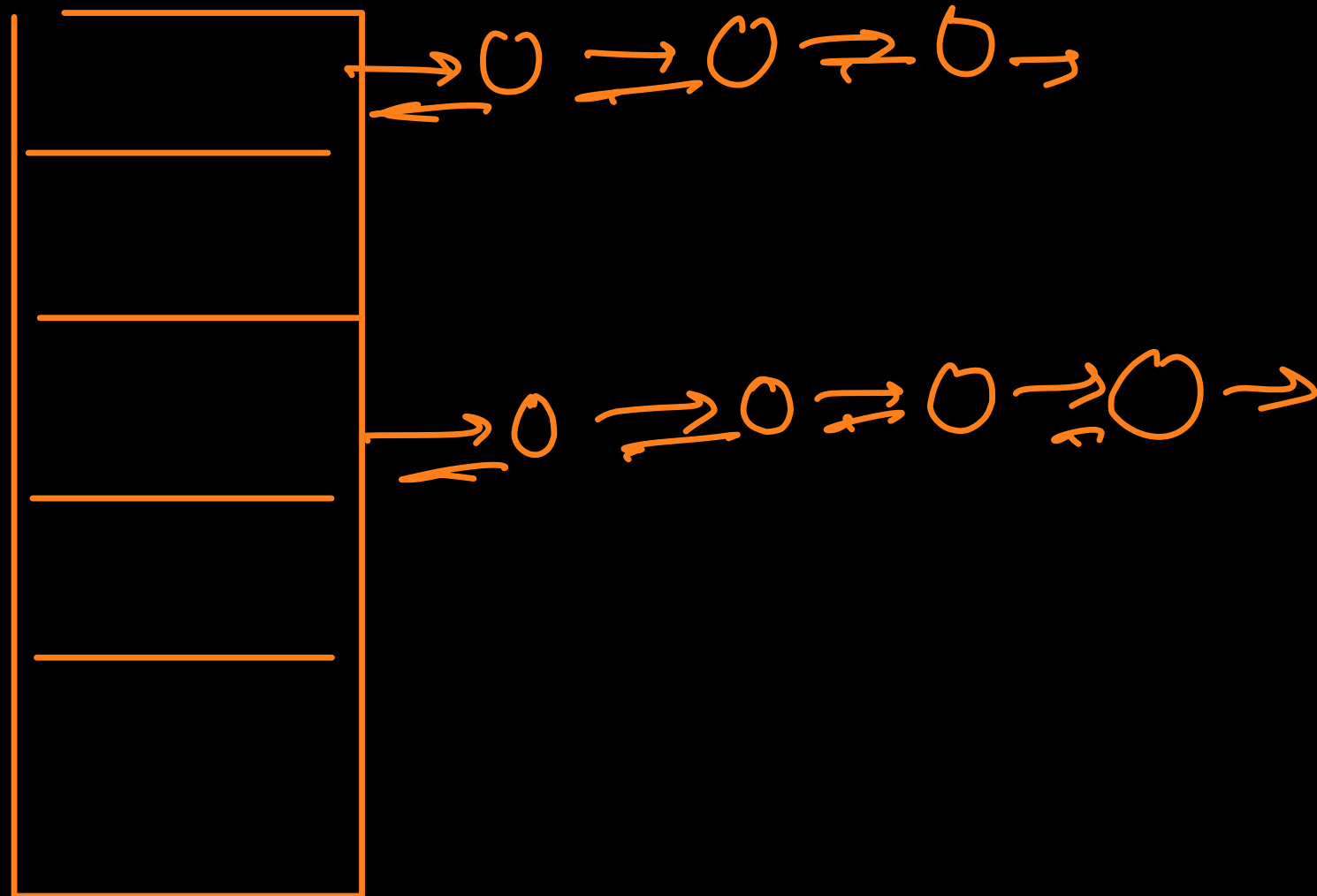
```
}
```

\rightarrow

$\rightarrow \underline{\underline{O(n)}}$

$\frac{n \times 1}{n}$

LL



find $O(n)$

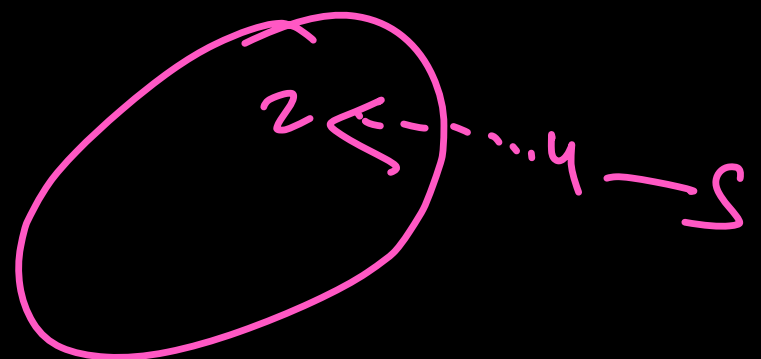
insert $O(1)$

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

$$\frac{n \times n}{n} \rightarrow \underline{\underline{O(n)}}$$

union(0,1)
 union(2,1) ←
 union(3,2)
 union(4,3)
n

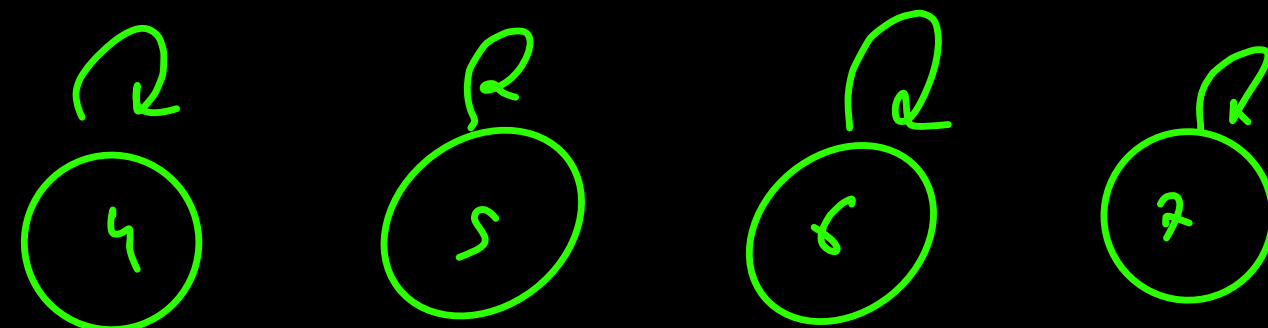
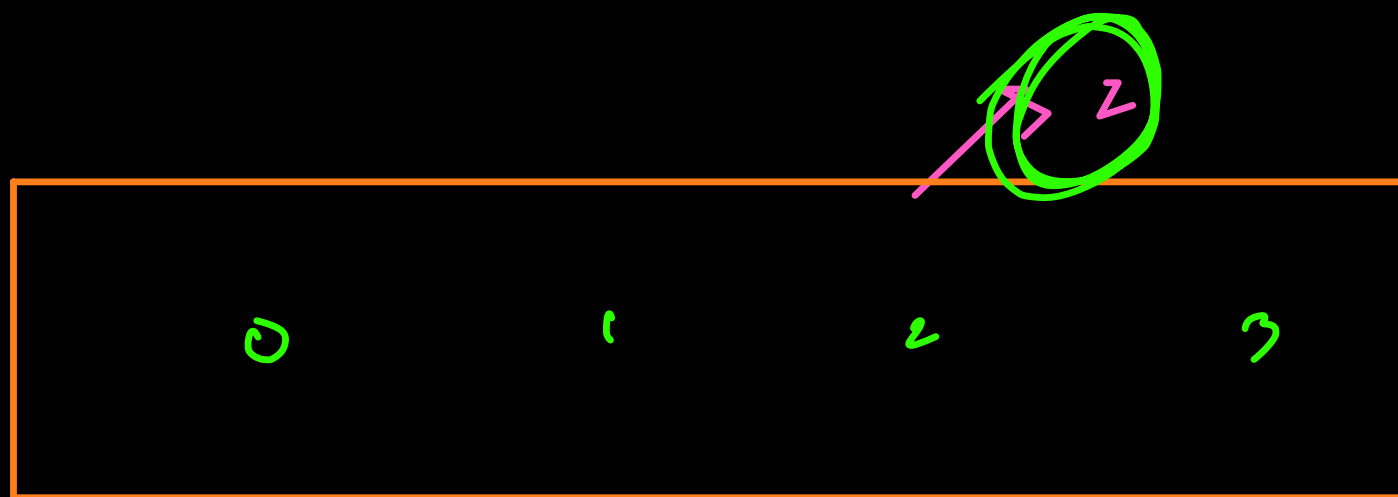
n union → n



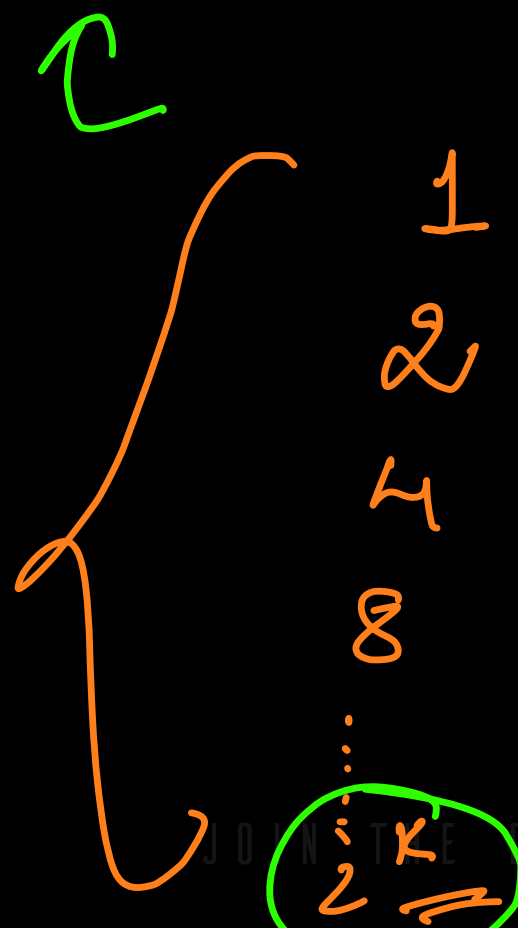
Union By Size \rightarrow smaller \rightarrow larger

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

Size \rightarrow
 $size[1] = + size[2]$



union (0,1)
 union (2,3)
 union (1,2)



$k \approx \log n$

$n \rightarrow \underline{\underline{\log n}}$

$$\begin{aligned}
 &\underline{\underline{2^k \leq n}} \rightarrow \underline{\underline{k \leq \log n}} \\
 &\quad \quad \quad \underline{n \times \log n} \rightarrow \underline{\underline{O(\log n)}} \\
 &\quad \quad \quad \underline{n_{\text{union}}}
 \end{aligned}$$

Qⁿ if for an element we need to find parent,
 what is one of the good D.S. to represent
 child & parent ?? \rightarrow Tree

0 1 2 3 4 5 6 7

1	3	3	3	5	3	6	7
---	---	---	---	---	---	---	---

→ par

1	2	1	6	1	2	1	1
---	---	---	---	---	---	---	---

→ size

union(0,1)

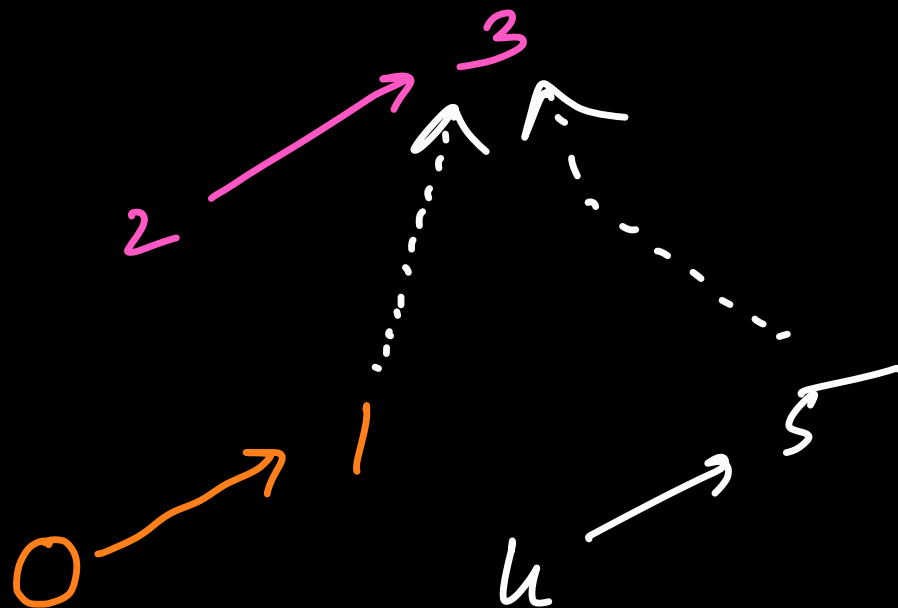
union(2,3)

union(2,0)

find(0)

un(4,5)

un(0,4)



union by Rank

void union (a, b) {

 a = find(a)

 b = find(b)

 if (sz[b] < sz[a]) {

 sz[a] += sz[b]

 par[b] = a

 } else {

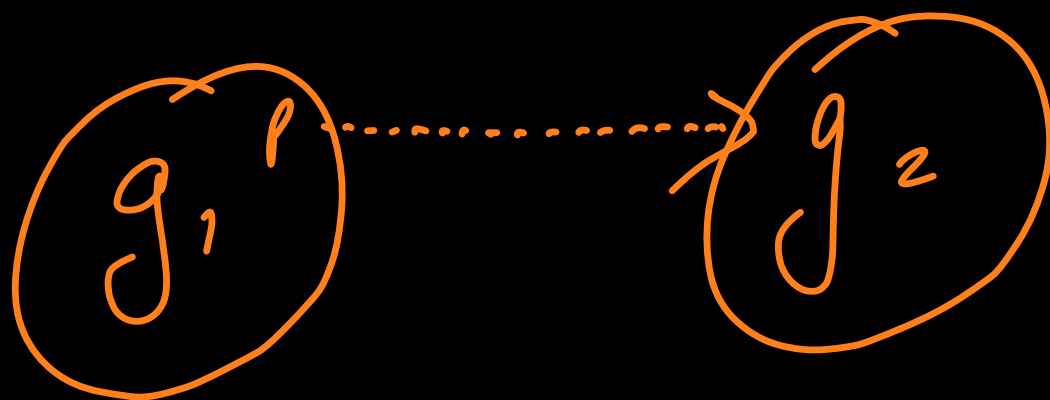
 sz[b] += sz[a]

 par[a] = b

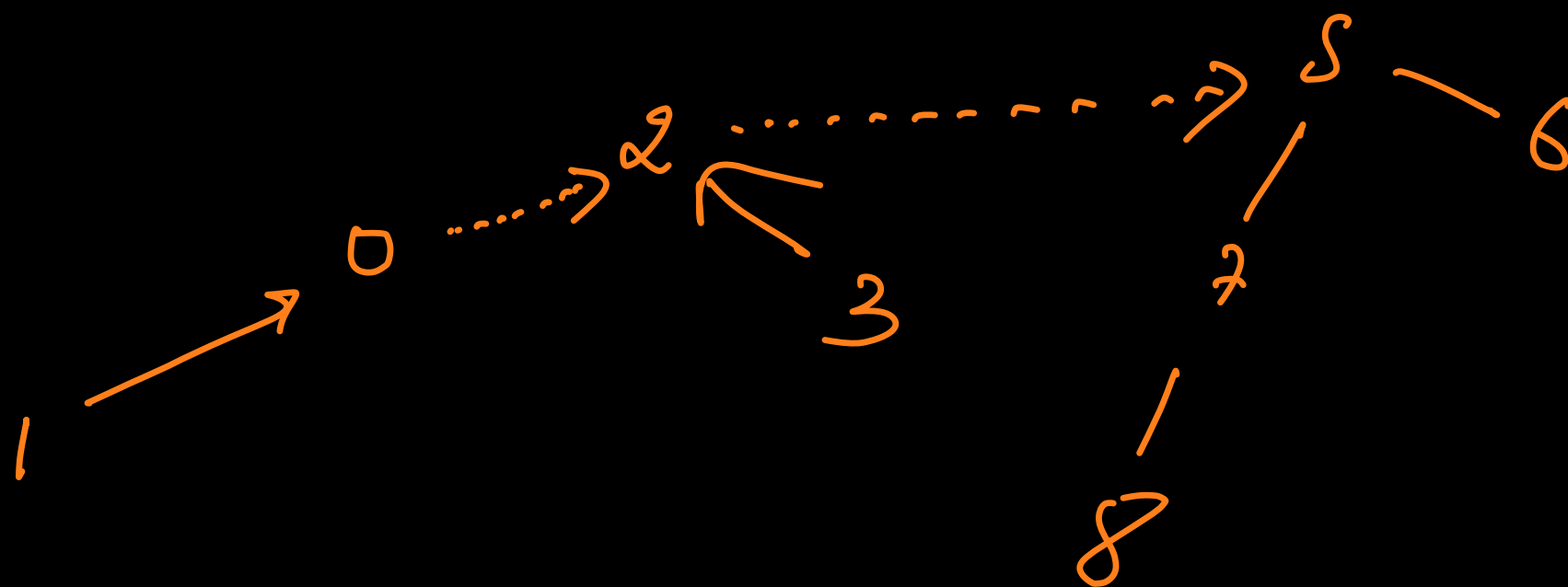
}

}

```
int find(x) {  
    if (par[x] == x) return x;  
    return find(par[x]);  
}
```



$$S_2 g_1 < S_2 g_2$$

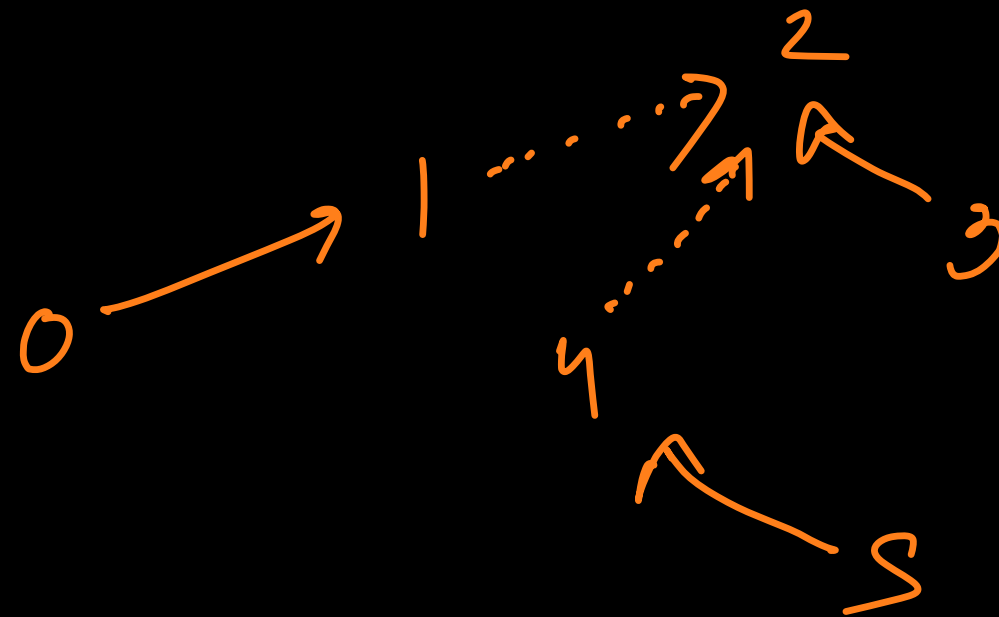


0	1	2	3	4	5	6	7
1	2	2	2	2	4	6	7

0	1	2	3	4	5	6	7
1	2	4	1	2	1	(1)	1

Rank:-

$\text{min}(0,1)$
 $\text{min}(2,3)$
 $\text{min}(1,3)$
 $\text{min}(4,5)$
 $\text{min}(2,4)$




```
void union(a, b) {
```

```
    a = find(a)
```

```
    b = find(b)
```

```
    if (rank[a] <= rank[b])
```

```
        par[a] = b
```

```
        rank[b]++
```

```
    } else {
```

```
        par[b] = a
```

```
        rank[a]++
```

```
    }
```

```
}
```

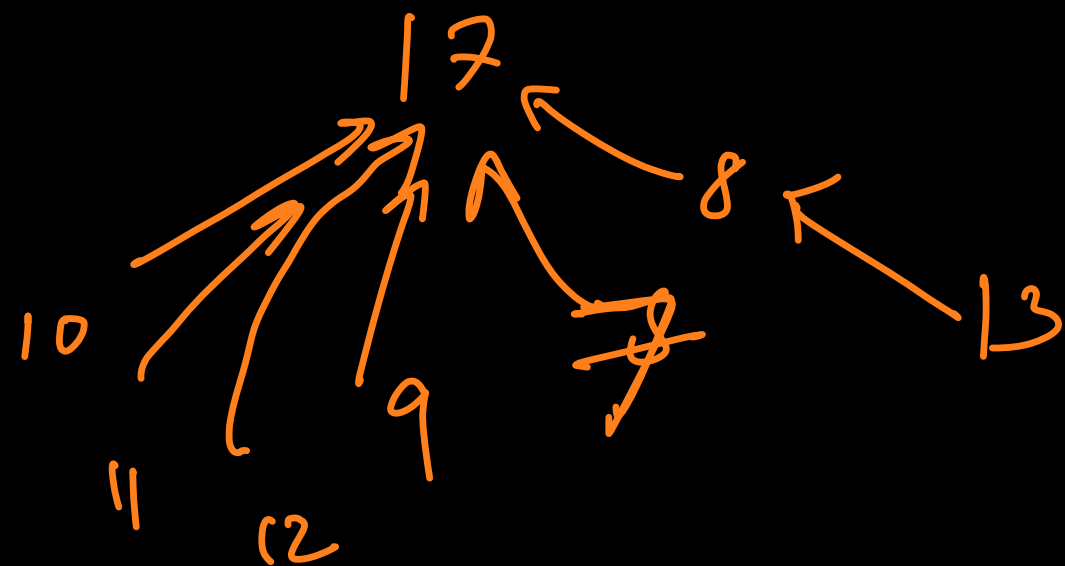
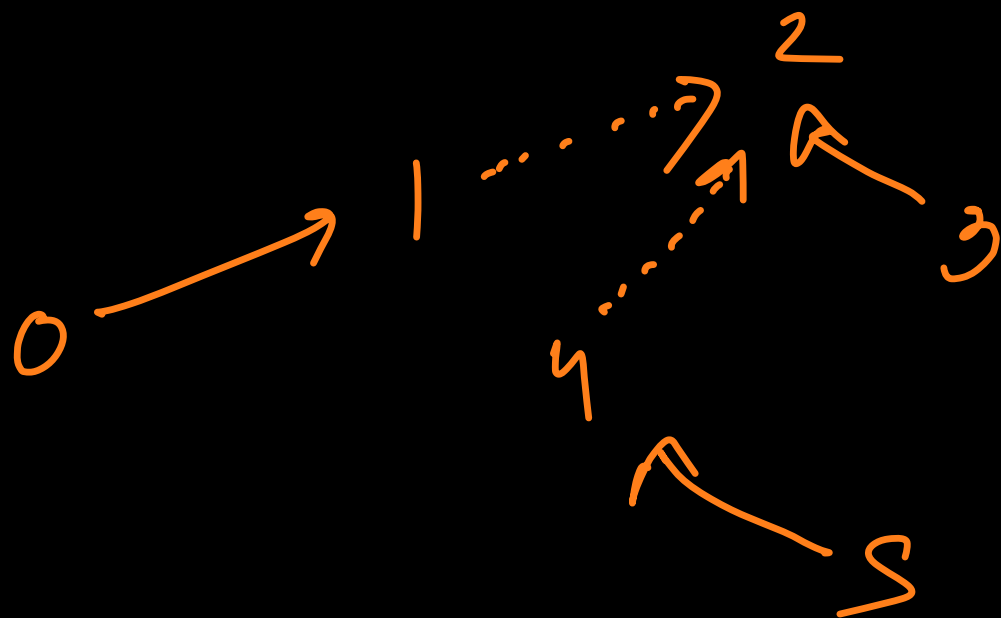
$O(\log n)$

union by size / union by rank

with path compression

$$\text{union}(S, T)$$

CC (16/17)



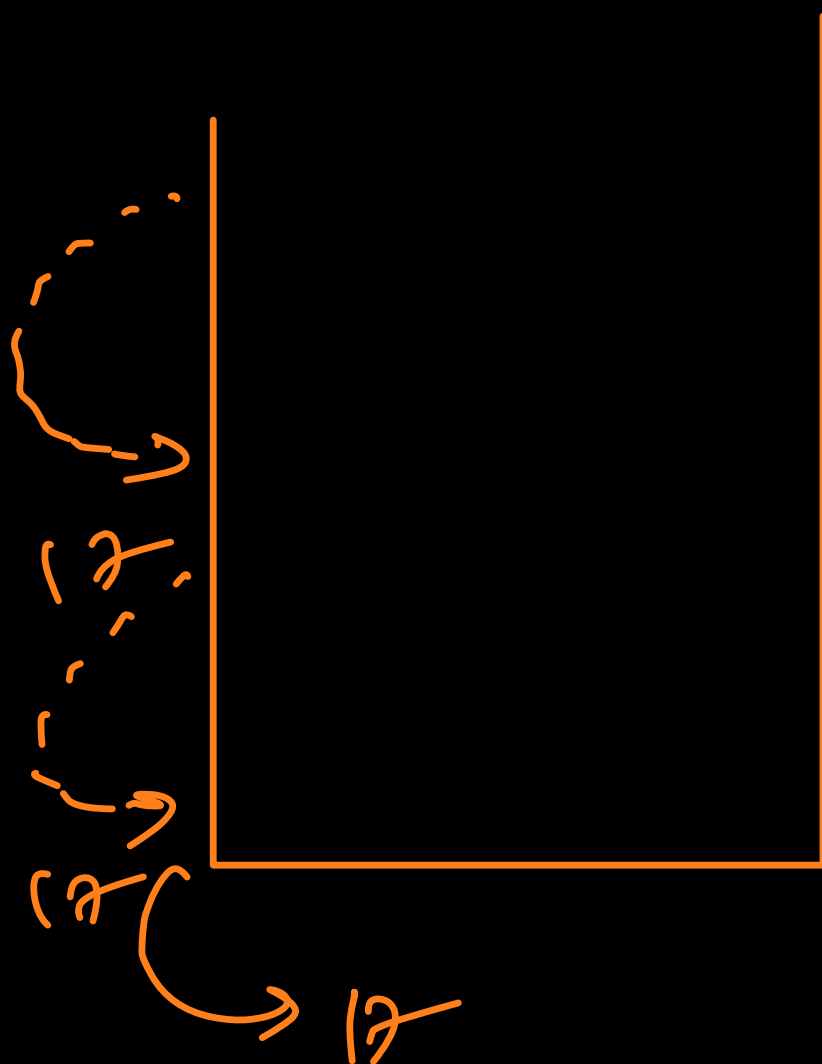
```

1) int find(x) {
2)   if (par[x] == x) return x;
3)   return par[x] = find(par[x]);

```

4) 3

			9	10	11	12		17	
...	17	17	17	17	...	17	...



inverse ackermann funcⁿ

$$n = \log_2 2^n$$

→ $\log^* n$ → this represents that if you

have a value n , and you repeatedly
apply $\log_2 n$ on this value then in how
many ops you can reduce it to ≤ 1

$$n = 65536$$

$$\hookrightarrow \log n = 16$$

$$\rightarrow \log_2 (65536) = 16$$

$$n = 65536$$

$$\underline{\underline{n = \log_2 7}}$$

??

$$\log_2 n \rightarrow \underline{\underline{25}}$$

$$\log_2^{(1)}(65536) \rightarrow \log_2^{(2)} 16 \rightarrow \log_2^{(3)} 4 \rightarrow \log_2^{(4)} 2$$

$$\rightarrow \underline{\underline{\log_2^{(1)}(1)}} \rightarrow \underline{\underline{0}}$$

$$n = 2^{32}$$

$$\log_{2^1} 2^{32}$$

(1)

$$\log_{2^2} 2^{32}$$

→

$$\log_{2^3} 2^{32}$$

→

$$\log_{2^5} 2^{32}$$

$$\approx \log_{2^2} 2^{32} \rightarrow \log_{2^1} 2^{32} \rightarrow 0$$

