

## 15-451/651 Algorithm Design & Analysis, Spring 2026

### Quiz 3 Solutions

---

**Problem:** The input is a graph with  $n$  vertices (and initially no edges) whose vertices are colored either **red** or **blue**. Give an algorithm that supports the following operations.

1. **ADDEDGE**( $u, v$ ): Add an undirected edge between vertices  $u$  and  $v$  in the graph,
2. **QUERY**(): Return the number of pairs of vertices  $u \neq v$  in the graph that are the same color and are in the same connected component (i.e., have a path between them).

**Desired runtime:** Over  $m$  operations, total time  $O((m+n)\log n)$ , equivalently amortized  $O(\log n)$  per operation.

---

We will utilize a union-find data structure. Each vertex maintains its parent in the tree. Additionally, the roots of each subtree contain the following information:

**Data at each root** Size of the subtree, as well as the number of red/blue vertices in its subtree. Call these  $\text{red}[v]$  and  $\text{blue}[v]$  respectively.

**ADDEDGE** Let's say the edge insertion is from  $u' \rightarrow v'$ , where  $u'$  is the root of  $u$ 's component and  $v'$  is the root of  $v$ 's component. Then the parent of  $u'$  is set to  $v'$ . We can update

$$\text{red}[v'] \leftarrow \text{red}[v'] + \text{red}[u'] \quad \text{and} \quad \text{blue}[v'] \leftarrow \text{blue}[v'] + \text{blue}[u'].$$

**QUERY.** We maintain a global counter of the number of pairs.

When the edge  $u' \rightarrow v'$  is added, we update this global counter by:

$$\binom{\text{red}[v'] + \text{red}[u']}{2} + \binom{\text{blue}[v'] + \text{blue}[u']}{2} - \left( \binom{\text{red}[v']}{2} + \binom{\text{blue}[v']}{2} \right) - \left( \binom{\text{red}[u']}{2} + \binom{\text{blue}[u']}{2} \right).$$

There are many other ways to do this, e.g., counting the number of “newly created pairs”, which is:

$$\text{red}[v']\text{red}[u'] + \text{blue}[v']\text{blue}[u'].$$

You can check that this is the same as above.

**Runtime** Updating the additional data take  $O(1)$  time, so the cost per operation is  $O(\log n)$  by many variants of union find.