University of Copenhagen
Applied High Performance Computing

Author: William Henrich Due

1 Definitions

Definition 1.1 (Union-Find)**.** The union-find data structure D represents a partition of a set S if:

1. $a \neq \emptyset$ for all $a \in \mathcal{C}(D)$
2. $a \cap b = \emptyset$ for all $a, b \in \mathcal{C}(D)$ where $a \neq b$
3. $\bigcup_{a \in \mathcal{C}(D)} a = S$

where $\mathcal{C} : \mathbb{P}(D) \to \mathbb{P}(\mathbb{P}(S))$ converts the data structure D into a set of sets of positive integers.

Definition 1.2. Union-find set The set $\mathbb{U}_n = \mathbb{P}(\mathbb{N}_{\leq n} \times (\mathbb{N}_{\leq n} \cup \{0\}) \times S)$ is the set of all union-find data structures of size n. Each element $U \in \mathbb{U}_n$ is a set of tuples (i, p, s) where:

1. $i \in \mathbb{N}_{\leq n}$ is a unique index identifying the element.
2. $p \in \mathbb{N}_{\leq n} \cup \{0\}$ is the index of the parent element. If $p = 0$ then the element is a root.
3. $s \in S$ is the value associated with the element.

Definition 1.3 (Cycle)**.** A cycle exists in a set $U \in \mathbb{U}_n$ if:

1. $i = p$ for some $(i, p, s) \in U$ or
2. $p_1 = i_2, p_2 = i_3, \ldots, p_m = i_1$ for some $(i_1, p_1, s_1), (i_2, p_2, s_2), \ldots, (i_m, p_m, s_m) \in U$ where $m > 1$.

Definition 1.4 (Find Root)**.** The root of an element i in a cycleless union-find data structure $U \in \mathbb{U}_n$ is defined as:

$$\mathcal{P}_U(i) = \begin{cases} i & \text{if } p = 0 \\ \mathcal{P}_U(p) & \text{if } p \neq 0 \end{cases} \text{ where } (j, p, s) \in U \wedge i = j.$$

Proposition 1.1 (Termination of Find Root)**.** The function $\mathcal{P}_U(i)$ defined in definition 1.3 terminates for all i where $1 \leq i \leq n$ if $U \in \mathbb{U}_n$ has no cycles.

Proof. The proof is trivially true. □

Proposition 1.2 (Representation of Union-Find)**.** The union-find structure $U \in \mathbb{U}_n$ fulfills definition 1.1 if:

1. Unique indices i.e. $|U| = |A|$ where $A = \{\pi_1(t) : t \in U\}$.
2. U has no cycles.
3. The following conversion function:

$$\mathcal{C}(D) := \{\{\pi_3(t) : t \in U \wedge \mathcal{P}_U(\pi_1(t)) = p\} : p \in \mathbb{N}_{\leq n}\} \setminus \{\emptyset\}$$

Proof. Let $U \in \mathbb{U}_n$ be a union-find data structure fulfilling the three criteria in proposition 1.2. We will show that U fulfills the three criteria in definition 1.1.

1. By definition of $\mathcal{C}$ it is clear that $a \neq \emptyset$ for all $a \in \mathcal{C}(D)$ since $\mathcal{C}$ only includes non-empty sets.
2. Let $a, b \in \mathcal{C}(D)$ where $a \neq b$.

$\square$