UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat d'Informàtica de Barcelona

FIB

Advanced Data Structures

Computer Science Department

Union Find Data Structure: An empirical analysis

Author:
Alex Herrero

Professor:
Conrado Marínez

February 17, 2025

Contents

List of Figures

List of Tables

1 Introduction

Given a binary relation R such that is:

- Reflexive: aRa
- Symmetric: $aRb \Rightarrow bRa$
- Transitive: $aRb \wedge bRc \Rightarrow aRc$

we say that R provides a partition Π of A into disjoint equivalence classes. That is, $\Pi = \{A_1, \ldots, A_k\}$ is defined as follows:

- $\forall A_i, A_j \in \Pi, A_i \cap A_j = \emptyset \iff A_i \neq A_j$
- $A = \bigcup_{i=1}^{k} A_i$
- $a \equiv b \iff a, b \in A_i$ for some $A_i \in \Pi$.

As every subset defines an equivalence class it is enough to represent every A_i with a single element $a \in A_i$ which is called the *representative* of A_i (as every other element of A_i is related to a by the properties of the binary relation previously defined).

Such idea was firstly introduced by Galler and Fisher[GF64] in Computer Science as the **union-find data structure** which is a data structure that stores partition of a set into disjoint sets. In particular **union find** consist of two main operation:

- Find Operation: Given two elements $a, b \in A$ determine if $\exists A_i \in \Pi$ s.t $a, b \in A_i$.
- Union Operation: Given two elements $a \in A_i$ and $b \in A_j$, *merge* A_i and A_j. That is, the result of this operation to the partition Π will be a new partition Π' such that $\Pi' = (\Pi \setminus \{A_i, A_j\}) \cup (A_i \cup A_j)$.

References

[GF64] Bernard A Galler and Michael J Fisher. An improved equivalence algorithm. *Communications of the ACM*, 7(5):301–303, 1964.