

Chapter 1

Disjoint set

Hlavným cieľom tejto kapitoly je vysvetliť, čo to disjoint set je, kde je možné túto datovú štruktúru použiť a akými spôsobmi ju možno implementovať.

1.1 Úvod

Set je datová štruktúra, ktorá združuje množinu vzájomne rozdielnych (disjunktných) dát. Rozdiel medzi setom a Disjoint setom (taktiež nazývaná Union-Find alebo Merge-find set) je ten, že Disjoint set rozdeľuje tieto prvky do podmnožín. Na počiatku máme množinu všetkých prvkov, ktoré sa nachádzajú v Disjoint sete a každý tento prvok tvorí samostatnú podmnožinu. Táto datová štruktúra definuje 3 operácie:

- **MakeSet** (x) - pridanie prvku x do množiny a jeho označenie ako samostatnej podmnožiny,
- **Find** (x) - zistí identifikátor množiny, v ktorej sa prvok x nachádza,
- **Union** (x, y) - zjednotenie množín, ktoré obsahujú prvky x a y .

Jednotlivé operácie tejto datovej štruktúry možno implementovať rôznym spôsobom – o týchto spôsoboch implementácie je možné sa dočítať nižšie.

1.2 Implementácia datovej štruktúry

Táto datová štruktúra je implementovaná ako les, kde každý strom reprezentuje jednu podmnožinu prvkov. Tieto stromy sa následne môžu zjednocovať a tvoriť tak novú podmnožinu.

Čo je to teda ten les a strom?

Pre definovanie lesu je potrebné najprv definovať čo je to strom, graf a ďalšie pojmy z teórie grafov:

- **(Neorientovaný) graf** - je usporiadaná dvojica (V, E) , kde:
 - V - je neprázdna množina vrcholov,
 - E - je množina hran (neusporiadaná dvojica vrcholov, značíme ju u, v),
- **Sled** - je sekvencia $v_0, e_1, v_1, e_2, v_2, \dots, e_k, v_k$, kde $e_i = \{v_{i-1}, v_i\}$ a $e_i \in E(G), \forall i \in 1, \dots, k$.
- **Cesta** - je sled, v ktorom sa neopakujú vrcholy.
- **Súvislý graf** - Graf G nazveme súvislým práve vtedy, keď pre každé 2 vrcholy u a v existuje $u - v$ - cesta,
- **Kružnica** - Majme $n \geq 3$, Kružnicou dĺžky n (kružnica pozostávajúca z n vrcholov) je graf $(\{1, \dots, n\}, \{\{i, i+1\} | i \in \{1, \dots, n-1\}\} \cup \{\{1, n\}\})$,
- **Podgraf** - Graf H je podgrafom grafu G práve vtedy, keď $V(H) \subseteq V(G)$
- **Acyklický graf** - Graf nazveme acyklický práve vtedy, keď neobsahuje ako podgraf kružnicu.
- **Strom** - Graf G nazveme stromom práve vtedy, keď je súvislý a zároveň acyklický,
- **Les** - Graf G nazveme lesom práve vtedy, keď je acyklický.

1.3 Použitie

Keďže už vieme, čo to ten Disjoint set je, bolo by dobré vedieť k čomu nám takáto datová štruktúra je a kde všade je možné ju použiť. Možné použitia:

- Udržovanie informácií o súvislých komponentách,
- Hľadanie minimálnej kostry na základe Kruskalovho algoritmu,
- Detekcia cyklu.