

Vyhledávání dat

- Více než 70% veškerého výpočetního času se spotřebuje na vyhledávání dat
- **Vyhledávání** spočívá ve zjištění, zda jistá hodnota **X** je obsažena v množině dat (datové struktuře) **M**.
 - výsledkem je obecně logická hodnota
 - často se za výsledek považuje místo (index, ukazatel), kde byla hodnota nalezena
- **Prvek množiny** – skládá se obecně ze dvou složek
 - **klíč**
 - pomocná hodnota sloužící k porovnávání prvků
 - pro klíč musí být definována relace **ekvivalence** a **uspořádání**
 - nejčastěji datový typ **integer**, **real**, **string**
 - **hodnota**
 - vlastní hodnota prvku
- **Vyhledávací prostor**
 - vhodná reprezentace množiny prvků
 - nejčastěji **pole**, **spojový seznam** nebo **binární strom**

Vlastnosti vyhledávacích metod

- **Sekvenčnost**
 - metoda přistupuje k prvkům vyhledávacího prostoru **sekvenčně** (postupné čtení) nebo **přímo** (prostřednictvím adresy)
- **Dynamičnost** – vyhledávací prostor se v průběhu zpracování nemění (**statický prostor**) nebo se mění vkládáním nebo rušením prvků (**dynamický prostor**)
- **Časová a prostorová složitost** – časová složitost operací a prostorová (paměťová) složitost reprezentace vyhledávacího prostoru
 - **paměťová složitost** je dána paměťovými nároky na reprezentaci zpracovávaných dat – nikoliv délkou programu
 - každá metoda vyhledávání vyžaduje **minimálně prostor pro n prvků**
 - některé metody navíc pomocné prostory
 - paměťová složitost vždy **lineárního řádu**, tj. $f(n) \approx k \cdot n$, $k \geq 1$ (konstanta)
 - výpočet **časové (operační) složitosti** bývá náročná úloha
 - zejména výpočet průměrné složitosti – je třeba odhadnout pravděpodobnost každé permutace vstupních dat
 - Obvykle se zjišťuje **asymptotická složitost pro nejhorší možný případ**
 - Operační složitost může sloužit ke klasifikaci tříd vyhledávacích metod