

## 20. Databáze - návrh

### 1. Princip databáze

- systém souborů s pevnou strukturou záznamu
- soubory jsou mezi sebou propojeny klíči
- součástí databáze je *systém pro řízení báze dat*
- organizovaná sbírka dat uložených a přístupných elektronickou formou
- cloudy, clusterová uložení

### 2. Druhy

- **Hierarchická:** jednotlivá data jsou uložena ve stromové struktuře, hierarchické schéma, vztah pouze  $1 : M$ , jeden z prvních modelů
- **Síťová:** zdokonalený model hierarchické dat. poskytuje vztah  $M : N$
- **Relační:** viz níže
- **Objektová:** jednotlivé informace jsou reprezentované ve formě objektů, OOP, existuje ještě hybridní objektově-relační databáze

### 3. Relační paradigma

- je založena na tabulkách, jednotlivé tabulky jsou propojeny (relace) klíči
- databázová tabulka je dvourozměrná struktura s záhlavím a tělem
  - slouce jsou atributy a řádky jsou záznamy
- každé atributy mají specifický datový typ a doménu (množina přípustných hodnot)
- pojem relační databáze souvisí s teorií množin, každá tabulka realizuje podmnožinu kartézského součinu množin přípustných hodnot všech sloupců - relací
  - kartézský součin:  $X \times Y = \{(x, y) : x \in X \wedge y \in Y\}$
- **kandidátní klíč:** atribut, který jednoznačně definuje záznam v tabulce, může se stát primárním klíčem, pakliže ne, potom hovoříme o alternativním klíči
- **primární klíč:** jedná se o primární identifikátor daného záznamu, může jím být jeden či více sloupců, nesmí být NULL, často se používají syntetické identifikátory - id
- **cizí klíč:** slouží pro vyjádření vztahů, relací, mezi databázovými tabulkami
- data uložená v tabulce by měla být integritní, musejí vyhovovat definovaným kritériím, zavádějí se tzv. *integritní omezení*
  - entitní (entita, neboli záznam, prvek z reálného světa, takže se kontroluje obecná integrita, třeba deduplicita), doménové (správnost datového typu)
- **stupeň vztahu:** unární (sám ze sebou, zaměstnanec je též nadřízený), binární (vztah mezi dvěma relacemi, slon je savec a savec je obratlovec), N-tární vztah (vztah mezi N relacemi, okurka je zelenina, jídlo, má zelenou barvu, ...)
- **kardinalita:** žádná, žádný vztah;  $1 : 1$  jeden záznam má relaci pouze s jedním; obdobně se dají odvodit  $1 : N$  (jedna kopírka několik zaměstnanců) a  $M : N$  (několik kopií a několik zaměstnanců)
- **parcialita:** neboli volitelnost vztahu, souvisí s kardinalitou

- **normální formy:** normalizace struktur databázových tabulek, navrhované tak, aby vykazovaly minimální redundanci

#### 4. Návrh diagramů

- způsob modelování dat, jejich reprezentace
- specifikace jazyka, který bude zpracovávat příkazy
- tolerance poruch, zabezpečení, ...
- planární grafy, kde vrcholy reprezentují jednotlivé tabulky, hrany potom reprezentují relace
- vztahové diagramy (název tabulku, její atributy, datové typy), spojení symbolizují vztahy mezi tabulkami