**SU Media:**

Student | Středník | ČWUT | AVC SH | Akropolis | iKariéra

# Datové modely

### <small>Z ČWUT</small>

**Obsah**

1 Konceptuální datový model, databázové modely dat
    1.1 Konceptuální modelování
    1.2 E-R Model
2 Fyzický pohled na data: záznam, soubor, tabulka, index
3 Techniky ukládání a přístupu k záznamům
4 B-stromy

# Konceptuální datový model, databázové modely dat

Databázový model je kolekce pojmů sloužících k modelování. Může být formální, matematická. Výsledkem databázového modelování je databázové schéma (popis struktury dat)

- schéma síťové databáze
- schéma relační databáze
- Schéma hierarchické - specíalní případ síťového. Schéma síťové databáze tvoří orientovaný graf. Pokud je tento graf stromem, hovoří se o hierarchických databázích.

## Konceptuální modelování

základní znalost jak interpretovat data je uložena v databázovém schématu. Konceptuální modely jsou pokusem umožnit vytvoření popisu dat v databázi nezávisle na fyzickém uložení databáze. Obvykle se řeší:

- struktura
- manipulace dat
- specifikace IO

čtyři základní principy konceptuálních modelů. (popisují objekty a jejich vzájemné vztahy)

- orientace na objekty - myslí se v objektech a nikoliv v identifikátorech, které je označují. Vyžadována složitá struktura objektů
- funkcionální podstata vztahů - do modelů se zahrnuje pojem atributu (E-R modely), dále funkce jako jediný základní konstrukt (DAPLEX, HIT), nebo systém funkcí (fragmenty v IFO)
- ISA hierarchie - umožňuje práci s nadtypy a podtypy objektů
- hierarchický mechanismus - pro konstrukci objektů z jiných objektů

## E-R Model

1. identifikuje typy entit jako třídy objektů stejného typu. ZAKAZNIK, FILM, KNIHA
2. identifikuje typy vztahů do kterých entity identifikovaných typů mohou vstupovat ZAKAZNIK (entita) MA_PUJCENOU (vztah) KNIHA(entita)
3. na základě přiměřené úrovně abstrakce přiřadí jednotlivým typům entit a vztahů atributy popisující blíže jejich vlastnosti JMENO (jméno, příjmení)
4. formuluje různá integritní omezení vyjadřující soulad s větší či menší přesností schématu s modelovanou realitou.

- Entita = je objekt reálného světa, který je schopen nezávislé existence a je jednoznačně odlišitelný od ostatních objektů. Konkrétní výskyt entitního typu.
- Vztah - je vazba mezi dvěma entitami
- Hodnota popisovaného typu - jednoduchý datový typ
- Atribut - funkce přiřazující entitám či vztahům hodnotu určující některou podstatnou vlastnost entity, nebo vztahu

Klíč = minimální množina atributů, jejichž hodnoty jednoznačně určují každou entitu

# Fyzický pohled na data: záznam, soubor, tabulka, index

- hromada = nehomogenní soubor záznamů obvykle proměnné délky. vkládané záznamy jsou umísťovány ve vymezeném prostoru disku fyzicky za sebou. Hledání O(N)
- sekvenční soubor = neuspořádaný obvykle jako halda. uspořádaný liší se, že je uspořádán dle klíče. Aktualizované záznamy se umisťují do neuspořádaného sekvenčního souboru (soubor aktualizací). na REORGANIZACI se setřídí do nové verze primárního souboru
- index sekvenční soubor = je tvořen primárním souborem (sekvenční setříděný soubor uspořádaný podle primárního klíče) s přídavnou víceúrovňovou strukturou index. úroveň primárního = 0. Další 1,2, ..
- indexovaný soubor = tvořen primárním souborem a indexy pro různé vyhledávací klíče. Neindexují se bloky, ale záznamy. Primární soubor tedy nemusí být setříděn, ani se nemusí nacházet v souvislé části vnější paměti.
- bitová mapa = jde o velmi efektivní možnost indexace řádků tabulek relační databáze. Prostě 1 na pozici v bitovém vektoru mi říká že k tomu patří i ten řádek, pokud 0 tak ne.
- přímý přístup = pomocí hasovací funkce se rozptyluje na d souborem obvykle bývá N=0.8M. jinak klasika
- B stromy

# Techniky ukládání a přístupu k záznamům

# B-stromy

B-strom řádu m nad vyhledávacím prostorem S = {p1, p2, ... pn } je kořenový strom těchto vlastností:

1. Každý uzel má maximálně 2m prvku p, vyjma kořene, který může mít minimálně 1 prvek.
2. Uzlem je buď list, který nemá žádné syny nebo vnitřní uzel s k+1 syny a k prvky pi, m <= k <= 2m. Jestliže p1, p2, ..pk jsou prvky a s0, s1, ... sk synové uzlu u, pak syn s0 obsahuje prvky menší než p1, syn si obsahuje prvky větší než pi a menší než pi+1
3. Všechny větve stromu mají stejnou délku

---

- Stránka byla naposledy editována v 17:46, 3. 2. 2007.