

PŘÍRODOVĚDECKÁ FAKULTA UNIVERZITY PALACKÉHO
KATEDRA INFORMATIKY

PŘEPISY PŘEDNÁŠEK

Databázové systémy 1

KMI/DATA1



Abstrakt

*Tento dokument obsahuje přepisy přednášek, které vedl doc. RNDr Vilém Vy-
chodil, Ph.D.. Uvedená práce (dílo) podléhá licenci Attribution-NonCommercial-
NoDerivs 3.0 Unported, a to včetně vložených souborů, to neplatí pro loga jiných
společností, jako je například logo PostgreSQL.*

Obsah

1	Přehled databázových systémů a jejich historie	7
1.1	Historie databázových systémů	7
1.1.1	Databáze založené na souborech	7
1.1.2	Databáze založené na síťovém modelu	8
1.1.3	Databáze založené na hierarchickém modelu	8
1.1.4	Databáze založené na objektovém modelu	8
1.1.5	Databáze založené na relačním modelu	8
2	Relační model dat	8
3	Struktura relačních dat	9
3.1	Požadavky na tabulky	9
	Seznam zkratk	10
	Seznam teorémů	11
	Bibliografie	12

Seznam obrázků

Seznam tabulek

1	Záhlaví tabulky	9
---	---------------------------	---

Seznam zdrojových kódů

1 Přehled databázových systémů a jejich historie

Databázový systém (anglicky [Database management system \(DBMS\)](#)) je soustava komplexního aplikačního vybavení, které je podloženo určitým teoretickým základem. Jeden bez druhého nemůže existovat (resp. může, což má ale za následek formální selhání [DBMS](#) jako takového), a tak je nutné znát obě strany pomyslné databázové barikády. Databázový systém tedy tvoří:

1. Aplikační software, který je obvykle používán jako rozhraní pro přístup k databázi samotné.
2. Teorie, která formálně podkládá návrh databáze, organizaci dat a obsahuje algoritmickou stránku věci.

Mějme na paměti, že obě částí databázových systémů se rozvíjely postupně a mnohdy metodou pokus - omyl. Obecně platí, že pokud selže teoretický základ, tak již ani sebelepší frontend nic nezmuže. V databázovém systému se objeví formální rozpor a jedinou cestou vpřed je začít znovu.

Hlavním úkolem [DBMS](#) je poskytovat *perzistentní uložení dat*, dále poskytnout svým uživatelům konzistentní rozhraní a případně nabídnout *transa-kční zpracování dat*. Nutno podotknout, že poslední bod nemusí být takovou samozřejmostí, jak by se mohlo zdát.

1.1 Historie databázových systémů

Potřeba organizace dat je stará jako lidstvo samo. Již staří Egypťané si vedli podrobné záznamy o výběru daní, stavbách chrámů a jiných činnostech. Zde můžeme hovořit o databázích založených na souborech.

1.1.1 Databáze založené na souborech

Souborem může být například papyrus, hliněná destička nebo (lépe) papír. Na papíře můžete být napsány v řádcích nějaké záznamy. Například seznam dlužníků nějakého podnikatele. *Vytvoření* takového seznamu je vskutku *lehké*. Představme si, že dlužník uprostřed seznamu splatil svůj dluh a bude ze seznamu vyškrtnut. Místo něj v seznamu vznikne prázdné místo. Zbytek seznamu se následně musí zkonsolidovat (přepsat na nový papír), aby vypadal konzistentně. Odtud plyne určitá *těžkopádnost* vyplývající z definice souboru a z určité *nízkoúrovňovosti* práce s ním. Přitom souborem může být myšlen i soubor na disku počítače.

Představme si navíc, že daný podnikatel si vede další soubor se seznamem, kde si u každého dlužníka značí jeho adresu, aby jej mohl v případě nutnosti navštívit. Jméno dlužníka tedy uchovává hned na dvou seznamech, přitom je přirozeně jasné, že stačí dlužníka evidovat jednou a poté se na něj „odkazovat.“ *Redundance dat* je tedy zřejmá.

1.1.2 Databáze založené na síťovém modelu

Průkopníkem této oblasti byl Charles Bachman. Aktuálně se toto paradigma považuje za dávno překonané, nicméně pro časy budoucí poskytlo několik dobrých poznatků. Typicky „síťový“ databázový systém je tvoří mj. i:

1. *Záznamy*, které mají určitý *typ*. Ten deklaruje, jaké *atributy* daný záznam obsahuje. Atributy mohou nabývat různých hodnot.
2. *Odkazy*, které reprezentují *vztahy* mezi jednotlivými záznamy.

Obecně se ví, že databázové systémy tohoto typu mohou poskytovat velmi rychlé prostředky pro získávání jednoduchých dat, ale na druhou stranu na komplexnější data se dotazuje hůře. Tvorba dat není rovněž jednoduchá, protože práce s odkazy požaduje vyšší míru sofistikace.

1.1.3 Databáze založené na hierarchickém modelu

Jedná se o speciální typ síťového modelu, kde bylo úmyslem tento model zjednodušit. Výsledkem bylo vytvoření hierarchie v záznamech a odkazech ve formě stromu¹.

1.1.4 Databáze založené na objektovém modelu

V objektovém modelu jsou hlavní prvky, jež reprezentují data, *objekty*, tedy serializované entity, které lze interpretovat pomocí nějakého vyššího programovacího jazyka. Prvním jazykem podporujícím takovéto pojetí byl Common Lisp.

1.1.5 Databáze založené na relačním modelu

Jedná se o aktuálně používaný model, který přináší určité výhody. Primárním nosičem informace v tomto modelu jsou *relace* (poněkud amatérsky je můžeme nazývat také tabulkami). Základy tohoto modelu položil Edgar Codd v roce 1970.

2 Relační model dat

Model má 3 základní komponenty:

1. *Struktury*, které uchovávají data a reprezentují výsledky dotazů.
2. *Integritní omezení*, která popisují vztahy mezi daty, které musí být splněny. Integritní omezení jsou v podstatě formule. Tato omezení slouží k popisu toho, jak mají data vypadat, slouží k odstranění jejich chyb.
3. *Manipulativní formalismy*, které říkají jak z uložených dat získat jiná data určitými operacemi.

Relační model je konkrétním modelem dat, avšak modely jako takové lze rozdělit do dvou skupin, tedy:

¹Strom je jednou ze základních grafových struktur. Jedná se o (neorientovaný) graf bez kružnic. Více napoví [1, str. 91 – 96].

Tabulka 1: Záhloví tabulky

ID	JMÉNO	ADRESA
15	Vyacheslav Drsoň	Peklo, 666
⋮	⋮	⋮

1. Abstraktní modely dat, které poskytují vyčerpávající logické definice datových struktur nebo operací s daty. Ty dohromady tvoří abstraktní formalismus. Tedy jakýsi abstraktní stroj, se kterým může uživatel operovat.
2. Konkrétní modely, pracující s persistentními daty.

3 Struktura relačních dat

Základním kamenem dat jsou tzv. „datové tabulky“, které obsahují záhlaví a další data.

1. Záhloví (viz. tabulka 1) je množina atributů a jejich typů. U tabulky se seznamem zaměstnanců může být atributem například jméno zaměstnance atp.
2. Vlastní data (neboli tělo) tabulky by u takovéto tabulky představovala data samotných zaměstnanců.

3.1 Požadavky na tabulky

1. Řádky by neměly mít žádné pořadí. Tabulka by měla být chápána jako množina řádků. A v množině na pořadí nezáleží.
2. To samé platí pro sloupce.
3. V definice množiny také vyplývá, že tabulka by neměla obsahovat identické řádky. Tento požadavek není v mnoha DBMS splněn.
4. Všechny atributy jsou regulární.

Tyto požadavky formuloval C. J. Date. Pokud tabulka splňuje body 1 až 5, pak je v 1. normální formě.

pokračovat

Seznam zkratk

DBMS Database management system

Seznam teorémů

Literatura

- [1] Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms*. English. Cambridge, MA: The MIT Press, 2000. ISBN: 0-262-03141-8.
- [2] Carlos Coronel, Steven Morris, and Peter Rob. *Database Systems: Design, Implementation, and Management*. English. 9th ed. Boston, MA: Course Technology, Nov. 23, 2009. ISBN: 0-538-74884-2.
- [3] Ramez A. Elmasri and Shamkant B. Navathe. *Fundamentals of Database Systems*. English. 6th ed. Boston, MA: Addison-Wesley, Apr. 9, 2010. ISBN: 0-136-08620-9.