Základy forenzních databází

Tereza Uhlíková

verze 2.0

Kdo jsem

Tereza Uhlíková

Ústav analytické chemie

skupina teoretické spektroskopie

místnost A235

https://web.vscht.cz/~uhlikovt/

tereza.uhlikova@vscht.cz

1. lekce

- Co je to databáze
- Proč, kdo, kdy, jak ...
- Něco málo z historie
- CAP problém

Dobrodružství kriminalistiky - televizní seriál
Dějiny psané Římem - Vojtěch Zamarovský

2. lekce

- Základy informatiky
- Software
  - informace
  - záznam informace
  - číselné soustavy
  - písmenné kódy
- Hardware
  - Alan Turing, John von Neumann a počítač
  - historie vývoje počítače & super počítač
  - procesor a datová uložiště
- Architektura databází

Alan Turing - Enigma
Contact film z roku 1997; seti@home

3. lekce

- Algoritmus
  - vlastnosti
  - zápis
  - struktura
- Datové typy

Arthur C. Clarke - Devět miliard božích jmen

O čem budeme mluvit dnes

- Výroková logika
- Databáze
  - DB a SŘBD
  - databázové modely
- Relační model
  - klíče
  - kardinalita vztahu

Aghata Christie - Hercules Poirot

Logika

Logika

Boolean algebra - 1847 George Boole
1880 Charles Sanders Peirce - "The Simplest Mathematics"

Logika

Boolean algebra - 1847 George Boole
1880 Charles Sanders Peirce - "The Simplest Mathematics"

věda zkoumající vztah vyplývání (dedukce)
problematiku správného usuzování

Logika

Boolean algebra - 1847 George Boole
1880 Charles Sanders Peirce - "The Simplest Mathematics"

věda zkoumající vztah vyplývání (dedukce)
problematiku správného usuzování

úsudek:

Logika

Boolean algebra - 1847 George Boole
1880 Charles Sanders Peirce - "The Simplest Mathematics"

věda zkoumající vztah vyplývání (dedukce)
problematiku správného usuzování

úsudek:
z předpokladů (premis) vyplývá závěr (důsledek)
... *A1, A2, ..., An* implikuje *B*
pouze o formu úsudků, nikoliv o obsah
How to tell the truth without knowing what you are talking about

Detektiv

Přijdou tři logici do baru.

Detektiv

Přijdou tři logici do baru.
Číšník se ptá prvního “Dáte si všichni pivo?”

Detektiv

Přijdou tři logici do baru.
Číšník se ptá prvního “Dáte si všichni pivo?”
První logik:”Nevím.”

Detektiv

Přijdou tři logici do baru.
Číšník se ptá prvního “Dáte si všichni pivo?”
První logik:”Nevím.”
Druhý logik:”Nevím.”

Detektiv

Přijdou tři logici do baru.
Číšník se ptá prvního “Dáte si všichni pivo?”
První logik:”Nevím.”
Druhý logik:”Nevím.”
Třetí logik:”ANO”

Výroková logika

Výroková logika

Dvouhodnotová extenzionální logika

Výroková logika

Dvouhodnotová extenzionální logika

výroková
Jestliže bude pěkně a nebudu učit, půjdu ven. $p \wedge \neg q \Rightarrow r$

Výroková logika

Dvouhodnotová extenzionální logika

výroková
Jestliže bude pěkně a nebudu učit, půjdu ven. $p \wedge \neg q \Rightarrow r$

predikátová –
1. řádu *Není pravda, že všichni lidé jsou spokojení* $\neg\forall x$
2. řádu *Existuje vlastnost, kterou mají všichni lidé* $\exists P \forall x$

Výroková logika

Dvouhodnotová extenzionální logika

výroková
Jestliže bude pěkně a nebudu učit, půjdu ven. $p \wedge \neg q \Rightarrow r$

predikátová –
1. řádu *Není pravda, že všichni lidé jsou spokojení* $\neg\forall x$
2. řádu *Existuje vlastnost, kterou mají všichni lidé* $\exists P \forall x$

Paradox lháře "Tato věta je nepravdivá."

Základní operace

Základní operace

konjunkce $a \wedge b$
disjunkce $a \vee b$
negace $\neg a$

Základní operace

konjunkce $a \wedge b$
disjunkce $a \vee b$
negace $\neg a$
slovy například:
konjugace: *Svítí slunce a zároveň prší.*
disjunkce: *Svítí slunce nebo prší.*
negace: *Nesvítí slunce.*

Z těchto základních operací lze odvodit všechny ostatní.
Existují zákony, dle kterých lze jednotlivé kombinace výroků převádět na jiné. Na příklad
De Morganovy zákony:
NAND $\neg(a \wedge b) = \neg a \vee \neg b$
NOR $\neg(a \vee b) = \neg a \wedge \neg b$

Základní operace

konjunkce $a \wedge b$
disjunkce $a \vee b$
negace $\neg a$
slovy například:
konjugace: *Svítí slunce a zároveň prší.*
disjunkce: *Svítí slunce nebo prší.*
negace: *Nesvítí slunce.*

Z těchto základních operací lze odvodit všechny ostatní.
Existují zákony, dle kterých lze jednotlivé kombinace výroků převádět na jiné. Na příklad
De Morganovy zákony:
NAND $\neg(a \wedge b) = \neg a \vee \neg b$
NOR $\neg(a \vee b) = \neg a \wedge \neg b$
slovy: *Není pravda, že svítí slunce a zároveň prší.* Je stejné jako: *Nesvítí slunce nebo neprší.*
Není pravda, že svítí slunce nebo prší. Je stejné jako: *Nesvítí slunce a zároveň neprší.*

Základní operace

implikace $a \rightarrow b = \neg a \vee b$

Základní operace

implikace $a \rightarrow b = \neg a \vee b$

Pokud je v IČ spektru silný pás u $3500\,\mathrm{cm}^{-1}$ *pak molekula obsahuje OH skupinu.*

To je stejné jako: *Není v IČ spektru silný pás u 3500 nebo molekula obsahuje OH skupinu.*

ekvivalence $a \leftrightarrow b = (a \wedge b) \vee (\neg a \wedge \neg b) = (a \vee \neg b) \wedge (\neg a \vee b)$

Základní operace

implikace $a \rightarrow b = \neg a \vee b$
Pokud je v IČ spektru silný pás u $3500\,\mathrm{cm}^{-1}$ *pak molekula obsahuje OH skupinu.*
To je stejné jako: *Není v IČ spektru silný pás u 3500 nebo molekula obsahuje OH skupinu.*

ekvivalence $a \leftrightarrow b = (a \wedge b) \vee (\neg a \wedge \neg b) = (a \vee \neg b) \wedge (\neg a \vee b)$
Existuje v IČ spektru silný pás u $3500\,\mathrm{cm}^{-1}$ je ekvivalentní tomu, že *látka obsahuje OH skupinu.*
Nebo *Existuje pás u* $3500\,\mathrm{cm}^{-1}$ *a zároveň obsahuje OH skupinu nebo neexistuje pás u* $3500\,\mathrm{cm}^{-1}$ *a zároveň neobsahuje OH skupinu.*
Lze to však napsat i jinak: *Existuje pás u* $3500\,\mathrm{cm}^{-1}$ *nebo neobsahuje OH skupinu a zároveň neexistuje pás u* $3500\,\mathrm{cm}^{-1}$ *nebo obsahuje OH skupinu.*

Ověření toho, že jsou si tyto operace rovny, se dělá pomocí pravdivostních tabulek.

Pravdivostní tabulky

Pravdivostní tabulky

a	b	$a \wedge b$	$a \vee b$	$\neg a$	$a \rightarrow b$	$a \leftrightarrow b$	$a \oplus b$
0	0	0	0	1	1	1	0
1	0	0	1	0	0	0	1
0	1	0	1		1	0	1
1	1	1	1		1	1	0

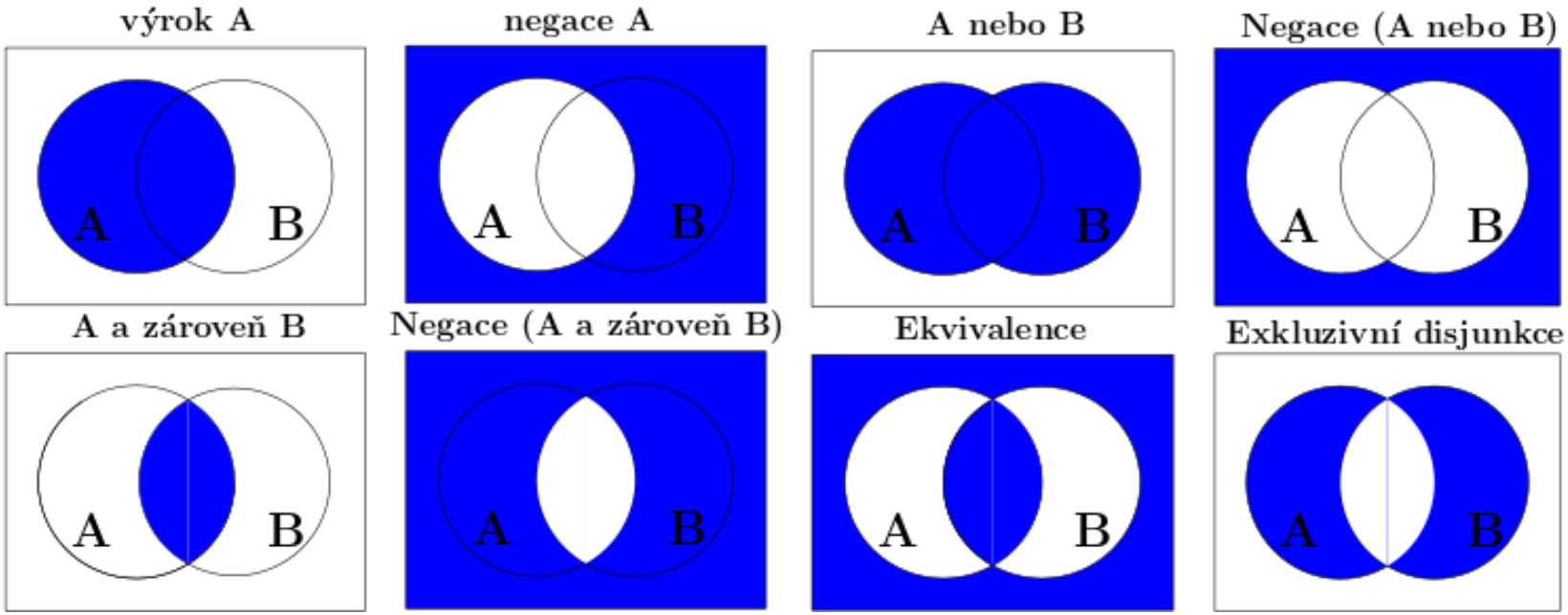
Exkluzivní disjunkce (XOR)

exkluzivní disjunkce
$a \oplus b = \neg(a \leftrightarrow b) = (a \vee b) \wedge \neg(a \wedge b) = (a \vee b) \wedge (\neg a \vee \neg b) =$
$(a \wedge \neg b) \vee (\neg a \wedge b)$
opak ekvivalence tedy neekvivalence

šifrování pomocí funkce XOR

Vstupní text:	0111011010101
Klíč:	1011000100100
$a \leftrightarrow b$	0011100001110
Výsledek XOR:	1100011110001

Vennovy diagramy



Tautologie

Tautologie

Logické pravdy; pravda při jakékoli interpretaci

Tautologie

Logické pravdy; pravda při jakékoli interpretaci

$a \leftrightarrow \neg\neg a$ zákon dvojité negace

Tautologie

Logické pravdy; pravda při jakékoli interpretaci

$a \leftrightarrow \neg\neg a$ zákon dvojité negace

$\neg(a \wedge \neg a)$ zákon sporu

Tautologie

Logické pravdy; pravda při jakékoli interpretaci

$a \leftrightarrow \neg\neg a$ zákon dvojité negace

$\neg(a \wedge \neg a)$ zákon sporu

Není pravda, že jsem v B14 a zároveň nejsem v B14.

$a \vee \neg a$ zákon vyloučeného třetího

Tautologie

Logické pravdy; pravda při jakékoli interpretaci

$a \leftrightarrow \neg\neg a$ zákon dvojité negace

$\neg(a \wedge \neg a)$ zákon sporu

Není pravda, že jsem v B14 a zároveň nejsem v B14.

$a \vee \neg a$ zákon vyloučeného třetího

Jsem v B14 nebo nejsem v B14. Nikde jinde být nemohu, protože mám jen dvě možnosti.

De Morganovy zákony jsou tautologií.

$\neg(a \wedge b) \leftrightarrow (\neg a \vee \neg b)$ negovaná konjunkce je disjunkcí negací

Tautologie

Logické pravdy; pravda při jakékoli interpretaci

$a \leftrightarrow \neg\neg a$ zákon dvojité negace

$\neg(a \wedge \neg a)$ zákon sporu

Není pravda, že jsem v B14 a zároveň nejsem v B14.

$a \vee \neg a$ zákon vyloučeného třetího

Jsem v B14 nebo nejsem v B14. Nikde jinde být nemohu, protože mám jen dvě možnosti.

De Morganovy zákony jsou tautologií.

$\neg(a \wedge b) \leftrightarrow (\neg a \vee \neg b)$ negovaná konjunkce je disjunkcí negací

$\neg(a \vee b) \leftrightarrow (\neg a \wedge \neg b)$ negovaná disjunkce je konjunkcí negací

Tautologie

Logické pravdy; pravda při jakékoli interpretaci

$a \leftrightarrow \neg\neg a$ zákon dvojité negace

$\neg(a \wedge \neg a)$ zákon sporu

Není pravda, že jsem v B14 a zároveň nejsem v B14.

$a \vee \neg a$ zákon vyloučeného třetího

Jsem v B14 nebo nejsem v B14. Nikde jinde být nemohu, protože mám jen dvě možnosti.

De Morganovy zákony jsou tautologií.

$\neg(a \wedge b) \leftrightarrow (\neg a \vee \neg b)$ negovaná konjunkce je disjunkcí negací

$\neg(a \vee b) \leftrightarrow (\neg a \wedge \neg b)$ negovaná disjunkce je konjunkcí negací

$\neg(a \rightarrow b) \leftrightarrow (a \wedge \neg b)$ převod implikace na konjunkci

Tautologie

Logické pravdy; pravda při jakékoli interpretaci

$a \leftrightarrow \neg\neg a$ zákon dvojité negace

$\neg(a \wedge \neg a)$ zákon sporu

Není pravda, že jsem v B14 a zároveň nejsem v B14.

$a \vee \neg a$ zákon vyloučeného třetího

Jsem v B14 nebo nejsem v B14. Nikde jinde být nemohu, protože mám jen dvě možnosti.

De Morganovy zákony jsou tautologií.

$\neg(a \wedge b) \leftrightarrow (\neg a \vee \neg b)$ negovaná konjunkce je disjunkcí negací

$\neg(a \vee b) \leftrightarrow (\neg a \wedge \neg b)$ negovaná disjunkce je konjunkcí negací

$\neg(a \rightarrow b) \leftrightarrow (a \wedge \neg b)$ převod implikace na konjunkci

Není pravda, že pokud bude pěkně půjdu ven. Je ekvivalentní k nebude pěkně a zároveň nepůjdu ven.

$(a \leftrightarrow b) \leftrightarrow (a \rightarrow b) \wedge (a \rightarrow b)$ ekvivalence je obousměrná implikace

Tautologie

Logické pravdy; pravda při jakékoli interpretaci

$a \leftrightarrow \neg\neg a$ zákon dvojité negace

$\neg(a \wedge \neg a)$ zákon sporu

Není pravda, že jsem v B14 a zároveň nejsem v B14.

$a \vee \neg a$ zákon vyloučeného třetího

Jsem v B14 nebo nejsem v B14. Nikde jinde být nemohu, protože mám jen dvě možnosti.

De Morganovy zákony jsou tautologií.

$\neg(a \wedge b) \leftrightarrow (\neg a \vee \neg b)$ negovaná konjunkce je disjunkcí negací

$\neg(a \vee b) \leftrightarrow (\neg a \wedge \neg b)$ negovaná disjunkce je konjunkcí negací

$\neg(a \rightarrow b) \leftrightarrow (a \wedge \neg b)$ převod implikace na konjunkci

Není pravda, že pokud bude pěkně půjdu ven. Je ekvivalentní k nebude pěkně a zároveň nepůjdu ven.

$(a \leftrightarrow b) \leftrightarrow (a \rightarrow b) \wedge (a \rightarrow b)$ ekvivalence je obousměrná implikace

Je pěkně je ekvivalentní s tím, že jsem venku. Bude-li pěkně budu venku a zároveň budu-li venku bude pěkně.

Pravdivostní tabulky

Opět si můžeme ověřit pomocí pravdivostních tabulek.

a	b	$\neg(a \wedge b)$	$\neg a \vee \neg b$	$\neg(a \wedge b) \leftrightarrow (\neg a \vee \neg b)$
0	0	1	1	1
1	0	1	1	1
0	1	1	1	1
1	1	0	0	1

Příklad

Určete, kdo je nevinný, když bylo zjištěno, že:
1) Z byl na místě činu právě tehdy, když tam nebyl ani jeden z dvojice X, Y.
2) Na místě činu nebyl podezřelý Z nebo není pravda, že tam byl alespoň jeden z dvojice X, Z.
3) Jestliže není pravda, že na místě činu byl X s Y, pak tam byl Z.

Úvod do databáze

Databáze neboli datová základna (Data Base)
je místo, kam se ukládají určitým způsobem organizované a strukturované údaje.

= data i software

Úvod do databáze

Databáze neboli datová základna (Data Base)
je místo, kam se ukládají určitým způsobem organizované a strukturované údaje.

= data i software

DBMS - DataBase Management System
SŘBD - Systém Řízení Báze Dat
je softwarové vybavení, které zajišťuje práci s databází.

Úvod do databáze

Databáze neboli datová základna (Data Base)
je místo, kam se ukládají určitým způsobem organizované a strukturované údaje.

= data i software

DBMS - DataBase Management System
SŘBD - Systém Řízení Báze Dat
je softwarové vybavení, které zajišťuje práci s databází.

- Definovat bázi

Úvod do databáze

Databáze neboli datová základna (Data Base)
je místo, kam se ukládají určitým způsobem organizované a strukturované údaje.

= data i software

DBMS - DataBase Management System
SŘBD - Systém Řízení Báze Dat
je softwarové vybavení, které zajišťuje práci s databází.

- Definovat bázi
- Konstruovat databázi

Úvod do databáze

Databáze neboli datová základna (Data Base)
je místo, kam se ukládají určitým způsobem organizované a strukturované údaje.

$$= \text{data i software}$$

DBMS - DataBase Management System
SŘBD - Systém Řízení Báze Dat
je softwarové vybavení, které zajišťuje práci s databází.

- Definovat bázi
- Konstruovat databázi
- Manipulovat s databází

Úvod do databáze

Databáze neboli datová základna (Data Base)
je místo, kam se ukládají určitým způsobem organizované a strukturované údaje.

= data i software

DBMS - DataBase Management System
SŘBD - Systém Řízení Báze Dat
je softwarové vybavení, které zajišťuje práci s databází.

- Definovat bázi
- Konstruovat databázi
- Manipulovat s databází

databázový systém x programovací jazyk

Příklady DBMS

Příklady DBMS

některých

- Microsoft Access
- MySQL
- SQLite
- PostgreSQL
- Oracle
- Informix
- DB2
- Sybase Adaptive Server Enterprise
- FileMaker
- Microsoft SQL Server
- Microsoft Visual FoxPro
- Progress
- CSQL
- OpenLink Virtuoso

Příklady programovacích jazyků

Příklady programovacích jazyků

některých

- Cobol
- `SQL`
- `Python`
- `C`
- `Pascal`
- `C++`
- `PHP`
- `Baltík`
- Basic
- Delphi
- PostScript
- Fortran
- Algol
- Matlab
- Perl

Data versus informace

Data versus informace

Data - konkrétní hodnoty (údaje), syntaktická složku

Informace - znalost, vyplývající z hodnot uložených v databázi, sémentická složka

Databázové modely

Databázové modely

- Lineární/sekvenční (textový soubor)
- Stromový (hierarchický model)
- Grafový (síťový model)
- Tabulkový (relační model)
- Objektový
- Objektově-relační
- XML - vychází z hierarchického,NoSQL databáze, generování pro html stránky

Hierarchický model

- stromová struktura
- vztahy podřízenosti a nadřízenosti
- každý nadřazený může mít více podřízených
- kažký podřízený má právě jednoho nadřazeného
- vhodné pro modelování typu část/celek

aplikace – rodokmen, evidence součástek v projektu Apollo

Síťový model

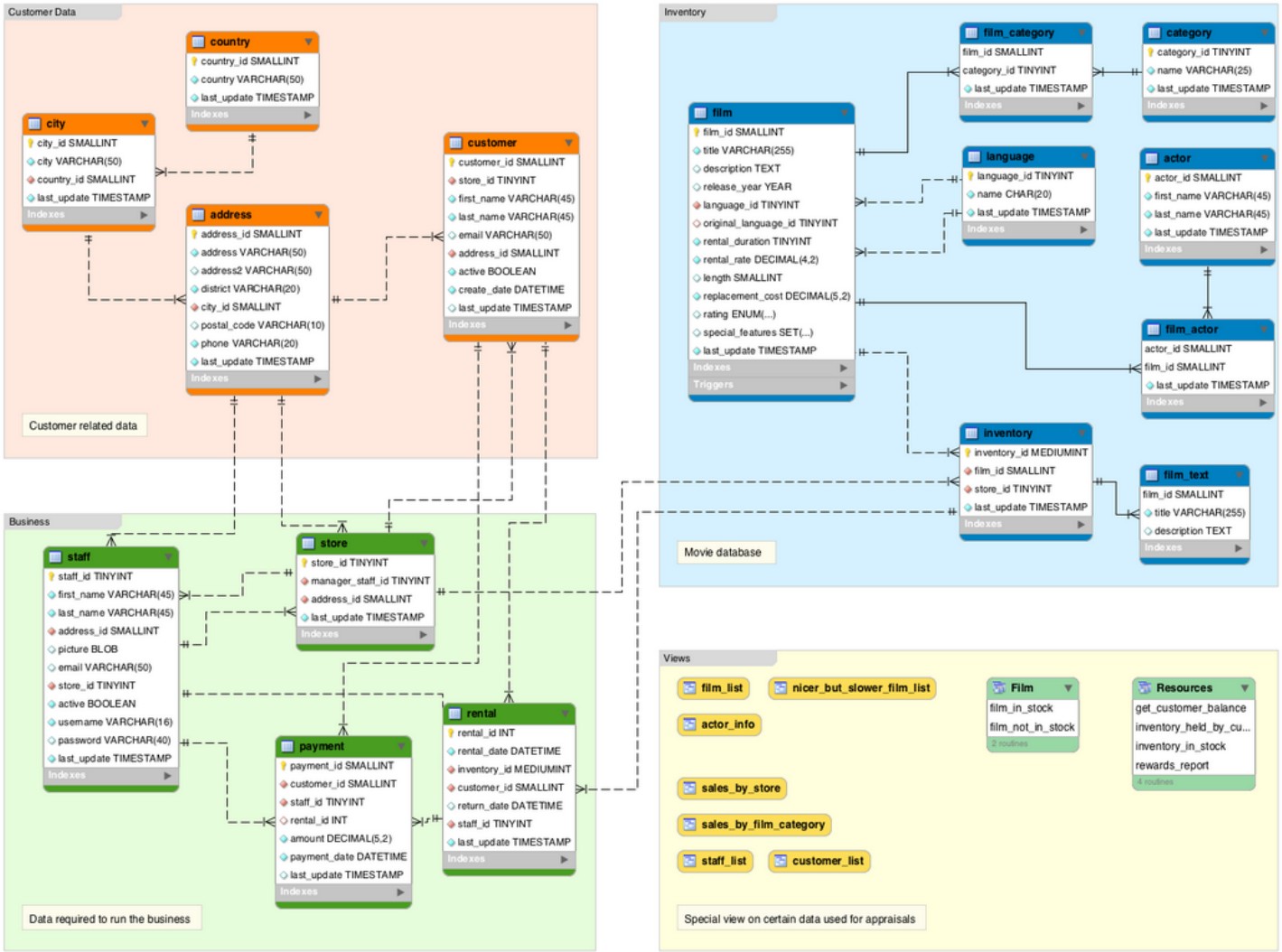
- orientovaný graf
- každý podřízený může mít více nadřízených

geografické informančí systémy

Relační model

- každá tabulka je přímo přístupná
- všechny tabulky mohou být navzájem propojené
- každá tabulka může mít více "nadřazených i podřazených"

Jak vypadá databáze



Objektový a relačně-objektový model

- třídimenzionální struktura
- objekty = data + metody.,.
- mezi objekty existuje skládání, dědění, závislost, klasifikace podle tříd, ... Strukturované informace není třeba rozdělovat jako v RDM
- protokol objektu je dán množinou přístupných zpráv (ne atributů jako v RMD)
- jedna množina (objektů) může s využitím polymorfismu obsahovat objekty s různou strukturou dat i metod
- je rozdíl mezi množinou objektů a třídou
- identita objektu je dána nejen vnitřními, ale i vnějšími vazbami, klíče jsou interní záležitostí

NÁVRH TABULEK

NÁVRH TABULEK

Relace, Tabulka, Entita

Relace, Tabulka, Entita

Co je to relace?
- spojeni, souvislost, vztah
- v tabulce jsou data v relaci mezi sloupcem a řádkem
=> tabulka
Objekt reality, který je zapsán pomocí jedné tabulky => entita

Relace ≈ Tabulka ≈ Entita

Tabulka obsahuje

Atribut: sloupec

- určují vlastnosti objektů (příjmení, fakulta,...)
- jedinečný název a - nepoužívat diakritiku a mezery, tedy např.: nazev_odberatele
- určený datový typ (číslo, text, logická hodnota, ...)

Záznam: řádek

- každý záznam je jednoznačně rozlišitelný - jeden řádek reprezentuje např. jednoho zaměstnance s hodnotami daných atributů
- mít svůj jedinečný identifikátor

id_zamestnance	prijmeni	jmeno	plat	datum_nastupu
1	Novák	Adam	32000	2013-05-02
2	Nová	Jana	45000	2013-06-11

Příklad

Navrhni tabulku pro lukostřelecké závody:
Dva lukostřelci Pavel a Michal byly na závodech. Pavel měl tři pokusy a stefil se postupně na čísla 4, 9 a 2.
Michal se trefil na čísla 1, 5 a 7.

Možné řešení

vícehodnotná položka

strelec	zasahy
Pavel	4, 9, 2
Michal	1, 5, 7

Identifikujte pole - při kolikátém zásahu se Pavel trefil nejblíže 10?

Upravené lepší řešení

vícehodnotná položka

špatně

strelec	zasahy
Pavel	4, 9, 2
Michal	1, 5, 7

správně

strelec	zasah1	zasah2	zasah3
Pavel	4	9	2
Michal	1	5	7

Návrh tabulek

vícehodnotná položka

špatně

strelec	zasahy
Pavel	4, 9, 2
Michal	1, 5, 7

správně

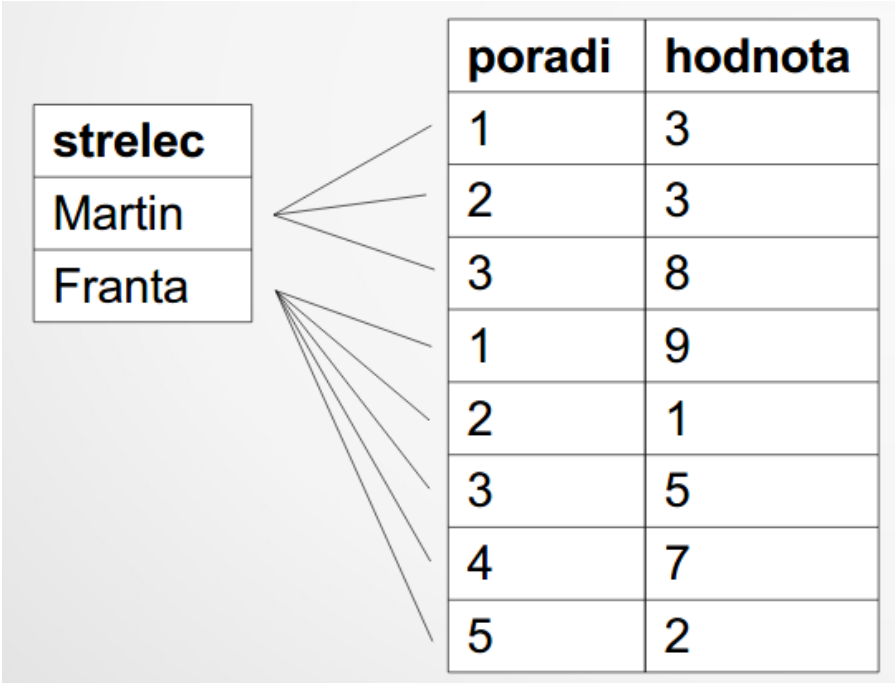
strelec	zasah1	zasah2	zasah3
Pavel	4	9	2
Michal	1	5	7

různý počet pokusů

strelec	z1	z2	z3	z4	z5	z6	z7
Martin	3	3	8				
Franta	9	1	5	7	2		

Dělení tabulek

Lze však tabulky rozdělit.



strelec
Martin
Franta

poradi	hodnota
1	3
2	3
3	8
1	9
2	1
3	5
4	7
5	2

Opakování jedné hodnoty

Lze tato tabulka zlepšit?

jmeno	mesto
Martin	Praha
Franta	Plzeň
Alena	Plzeň
Jana	Plzeň

Rozdělení tabulek

opakující se hodnoty
plýtvání místem
provede-li se změna, musí se opravit všude

špatně

jmeno	mesto
Martin	Praha
Franta	Plzeň
Alena	Plzeň
Jana	Plzeň

správně

jmeno
Martin
Franta
Alena
Jana

mesto
Praha
Plzeň

Určitelné informace

vypočítaná položka
– sloupec, který se dá vytvořit výpočtem z ostatních položek v záznamu, by neměl být součástí tabulky
– výpočet se provede v dotazu

špatně

obdelnik	sirka	vyska	obvod	obsah
malý	2	4	10	8
velký	30	15	90	450

Skripta, stránky a materiály

https://www.fi.muni.cz/~popel/lectures/bak_logika/
https://www.siliconrepublic.com/science/logic-maths-puzzle-george-boole
http://www.di.unito.it/~anselma/pdf/preprintBoole.pdf
Simultaneous Equations and Boolean Algebra in the Analysis of Judicial Decisions
Fred Kort Law and Contemporary Problems Vol. 28, No. 1, Jurimetrics (Winter, 1963), pp. 143-163 (21 pages)
https://en.wikipedia.org/wiki/Truth_table
https://sites.ff.cuni.cz/uisk/wp-content/uploads/sites/62/2016/01/Datov%C3%A9-modely-a-n%C3%A1vrhy-rela%C4%8Dn%C3%ADch-sch%C3%A9mat_Sk%C5%99ivan.pdf
http://www.databaze.chytrak.cz/modely.htm
http://books.fs.vsb.cz/dbacc20/Welcome.htm
http://books.fs.vsb.cz/
http://projekty.fs.vsb.cz/463/edubase/