

# DATABÁZOVÉ A ZNALOSTNÍ IS

P2

2008-02-28

## ENTITY-RELATION DIAGRAMS:

### Rozdíl mezi síťovým a datovým modelem:

- ✓ V síťovém modelu je odkaz na místo v paměti, cizí klíč, odkaz je skutečným odkazem do paměti.
- ✓ Relační datový model se chová podobně, ale máme-li hodnotu cizího klíče, zjistí se hodnota, ale ze samotné hodnoty nevíme, kde záznam leží, je nutné projít všechny záznamy a najít ho => bez použití indexování je databáze velmi pomalá.

### Objektový datový model:

- ✓ Podobný síťovému
- ✓ Je síťový, pokud doplníme o polymorfismus a další objektové funkce.

### Dva způsoby nákupu vánočního stromku:

- ✓ Seřadit šišky do řady, všechno pod 1,5 metru pryč, borovice krok vpřed atd.
- ✓ Obchod, v regálu šišky, vedle větvičky, v pytlích jehličí, na konci stonky, všichni chodí s lupou, tak si jí vezmeme taky a zjistíme, že na všem jsou čísla. Vybereme si tedy stonek, zapamatujeme číslo a dohledáme k tomu větvičky, jehličí a šišky. Sestavíme stromek dohromady a zjistíme, že se nám nelíbí, takže musíme totéž udělat znovu. Prodávač nám pak vybrané části stromku sesype do pytle a jdeme domů.

### Relační datový model – Relational Data Model:

- ✓ Values are atomic
- ✓ Values are scalar
- ✓ Values exist in domains, all values within a given domain are considered comparable
- ✓ Predicate logic is used to operate on entire tables as single units
- ✓ One column or group of columns is defined to be unique for each row, this is called the primary key and is used to distinguish rows from one to another
- ✓ A column or group of columns in a table may reference a column or group of columns in another table, this is the relation between foreign key and parent key
- ✓ Row and column subsets of a table can be defined, extracting a row subset is called “selection”, extracting a column subset is called “projection”, the result of these operations is also a table
- ✓ Multiple tables can be combined using set operations such as union, intersection, Cartesian product and difference, join is the operation of Cartesian product and selection together, the result of all these operations is also a table

### Rules of E.F. Codd for Relational Data Model:

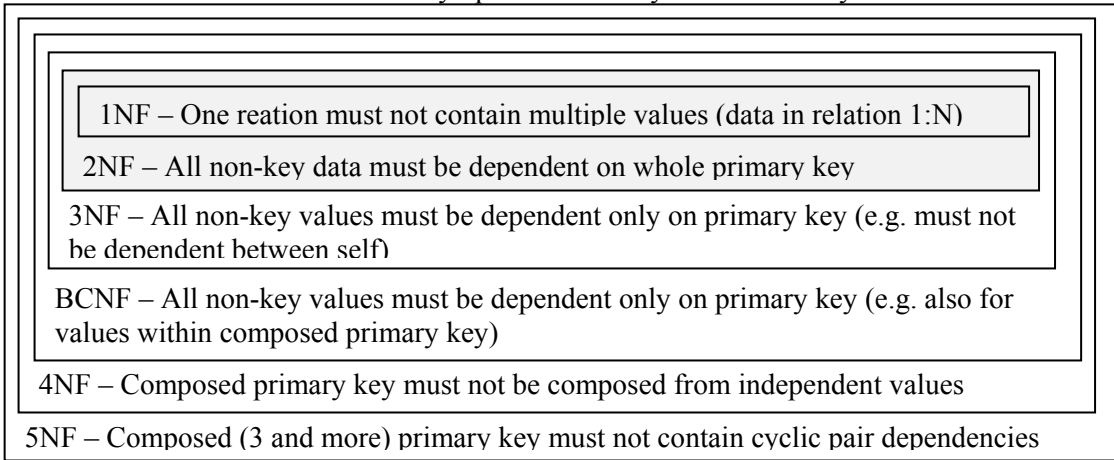
- ✓ The information rule
- ✓ The guaranteed access rule
- ✓ Systematic treatment of null values
- ✓ Active online catalog...
- ✓ ...?
- ✓ The view updating rule
- ✓ High-level insert, update, and delete
- ✓ Physical data independence
- ✓ Logical data independence
- ✓ Integrity independence
- ✓ Distribution independence
- ✓ The non-subversion rule

### Normalizace – Data Normalization

- ✓ Can be combined with other techniques – decomposition and synthesis
- ✓ Enables to find table structure with minimal redundancy and maximal consistency



- ✓ Is able to test how solution meets its requirements
- ✓ Non-deterministic approach – there is more “good” solutions
- ✓ Normalized data are not automatically optimal to real system effectively



- ✓ Square, curly
- ✓ 1NF:  $R(\underline{ABC}, \{A \rightarrow B, A \rightarrow C\})$   $R \Rightarrow R1(\underline{AC}) \oplus R2(\underline{BA})$
- ✓ 2NF:  $R(\underline{ABCD}, \{AB \rightarrow C, A \rightarrow D\})$   $R \Rightarrow R1(\underline{ABC}) \oplus R2(\underline{AD})$
- ✓ 3NF:  $R(\underline{ABC}, \{A \rightarrow B, A \rightarrow C, B \rightarrow C\})$   $R \Rightarrow R1(\underline{AB}) \oplus R2(\underline{BC})$
- ✓ BCNF:  $R(\underline{ABC}, \{C \rightarrow B, \dots\})$   $R \Rightarrow R1(\underline{AC}) \oplus R2(\underline{CB})$
- ✓ 4NF:  $R(\underline{ABC}, \{B \sim C, \dots\})$   $R \Rightarrow R1(\underline{AB}) \oplus R2(\underline{AC})$
- ✓ 5NF:  $R(\underline{ABCD}, \{ABC \twoheadrightarrow D, A \twoheadrightarrow B, B \twoheadrightarrow C, C \twoheadrightarrow A\})$   $R \Rightarrow R1(\underline{AB}) \oplus R2(\underline{BC}) \oplus R3(\underline{CA}) \oplus R3(\underline{ABCD})$