

Príklad 1 (20 b)

Veľký obchodný dom s nepretržitou prevádzkou si udržuje databázu:

Tovar(*IDT*, *názov*, *kategória*, *cena*)

Košík(*IDK*, *IDT*, *množstvo*)

Predaj(*IDK*, *dátum*).

Tabuľka *Košík* reprezentuje jednotlivé nákupy, tabuľka *Predaj* obsahuje údaj o čase, kedy sa nákup uskutočnil. Tabuľka *tovar* obsahuje údaje o tovaroch.

Kde *IDT* je identifikátor tovaru, *IDK* je identifikátor nákupu. Atribúty *názov*, *kategória*, *cena*, *množstvo*, *dátum* majú prirodzený význam. Navyše predpokladáme, že v databáze sú zabudované aritmetické predikáty $ADD(x, y, z) \Leftrightarrow z = x + y$ (môžete používať aj rozdiely napr. $x = z - y$), $MULT(x, y, z) \Leftrightarrow z = x * y$ (použiteľný aj pre podiely). A dátumový predikát *WEEKDAY*(dátum, deň), kde deň je reťazec (pondelok, útorok, ..., nedeľa). Tento predikát je pravdivý dátum je daný deň.

Sformulujte nasledovné dotazy v datalógu, kalkule a algebre na:

- Na tovary (*IDT*, *názov*, *cena*), ktoré sa predávajú každý deň týždňa.
- Dvojice tovarov, ktoré sa nikdy nevyskytli spolu v košíku.
- Dvojice tovarov, ktoré sa vyskytli súčasne v košíku a každý z predal sa vyskytuje v aspoň v 10% všetkých nákupov.

Príklad 2 (10b)

Daná je schéma $S = \{A, B, C, D, E, F, G, H\}$.

Platia závislosti: $F = \{EF \rightarrow ABCD, ACD \rightarrow G, AE \rightarrow F, BF \rightarrow C, C \rightarrow A, G \rightarrow CE\}$

Nájdite minimálne pokrytie, všetky kľúče a upravte schému **S** do 3NF nelámajúcej závislosti a BCNF. Nájdite 3NF nelámajúcu závislosti s najmenším počtom tabuliek.

Úloha 3 (10b)

- Definujte normálne formy 3NF, BCNF, 4NF a napíšte vzťah medzi nimi.
- Dokážte, že binárna relácia je vždy v BCNF.
- Čo to znamená, že dekompozícia schémy $R(x, y, z)$ na schémy $S(x, y)$ a $T(y, z)$ sa spája bezstrátovo?
- Čo je to trojschémová architektúra a k čomu slúži.
- Popíšte algoritmus testovania zachovania závislosti pri dekompozícii.

Úloha 4 (10b)

Popíšte spôsob vytvárania a využitie indexovo sekvenčných súborov. Vysvetlite k čomu slúži viac úrovní indexov. Čo je riedky a hustý index. Definujte B-strom. Popíšte algoritmus vloženia a vynechania prvku v B-strome. Porovnajete rôzne varianty B-stromov.

Príklad 5 (10b)

Je potrebné kódovať slová dĺžky 8 pozostávajúce z piatich znakov, ktorých pravdepodobnosti výskytu sú nasledovné: $\{a - 0.3, b - 0.25, c - 0.2, d - 0.15, e - 0.1\}$.

- Navrhните optimálnu metódu kompresie.
- Kompresujte slovo *accdbaaa*.
- Porovnajete s kompresiou Huffmanovým kódom.