Skúška E1 *20.1.2004*

Example 1 (20 points)
A warehouse maintain the database:

Product (IDT, name, category, price)
ShoppingBin (IDK, IDT, quantity)
Deal (IDK, date).

The table *ShoppingBin* records products and their quantities in a deal. The table *Deal* contains records about date of the deals. The table *Product* contains details about products.
IDT is an identifier of the product; *IDK* is an identifier of the deal. Attributes, *name, category, price, quantity* and *date* have their natural meaning. We assume that the database has built-in arithmetical predicates as $ADD(x, y, z) \Leftrightarrow z = x + y$ (it can be used for difference e.g. $x = z - y$), $MULT(x, y, z) \Leftrightarrow z = x * y$ (used for quotient, too), and date predicates as *WEEKDAY(date, weekday)* and *Month(date, month)*.

Please, formulate queries in datalog, calculus and algebra:

a) For products (IDT, name, price) sold but never sold in Monday.
b) Total price, separate 20% VAT included, for electronics (the category) sold in December.
c) Pairs of products sold but newer in a shopping bin together.

Example 2 (10 points)
Given a scheme **S** = { A, B, C, D, E, F, G }.
With dependencies: **F** = {EF → ACD, ACD → BEFG, AE → B, BF → C, C → A, G → EF}

Find a minimal cover, all the keys and the scheme **S** into 3.NF nonbreaking dependencies and BCNF. Try to avoid unnecessary decomposition.

Task 3 (10 points)

1. Define normal forms 3.NF, BCNF, 4.NF and write the relations among them.
2. Prove that each binary relation is in the BCNF.
3. What does it means, that the decomposition of the scheme ***R(x, y, z)*** in to shemes ***S(x, y)*** and ***T(y, z)*** joins loslessly?
4. What is it three scheme architecture and what one gains by it.
5. Describe an algorithm for testing dependencies preservation after decomposition.

Task 4 (10 points)
Describe schemes for dynamisation of the hashing (Larson, IBM, Litwin) and give some argument for linear expected complexities for operations (find, insert, delete).

Example 5 (10b)
We deal with compression of words of the length 8 consisting of five symbols with following probabilities { a – 0.3, b – 0.25, c – 0.2, d – 0.15, e – 0.1}.

a) Suggest an efficient compression method.
b) Compress the word bcdaaaaa.
c) Compare to Huffman coding.