Answers for Databases 2005 – Paper 5 Question 8 (tgg22)

(a) [Lecture 6 of syllabus.] The operators are (1) set-theoretic union, (2) set-theoretic intersection, (3) set-theoretic difference, (4) selection, and (5) projection. A full answer should briefly describe each operator.

(b) [Lectures 4 and 6 of syllabus.] Some of the possible differences are: the relational algebra is set-theoretic, while SQL is based on bags (multi-sets), the relational algebra is used for mathematical analysis and illustration, while SQL was desinged for usability, the relational algebra is a convention, not an international standard, the relational algebra is simple, while SQL is very complex. Some of the possible similarities are: both are query languages for relational databases, both are declarative, not imperative languages, both hide the details of data layout and implementation, and both return collections of tuples as answers.

(c) [Lectures 3 and 6 of the syllabus.] The relations $R(c, a) = \{\langle v, \ x \rangle\}$, $S(a, b) = \{\langle x, \ y \rangle\}$ is a counter-example (when $v \neq y$), since the left side of the equation is $\langle v, \ x, \ y \rangle$, while the right is $\{\}$.

(d) [Lectures 3, 8, and 9 of syllabus.] Normal Forms.

    (i) A relation R with FDs F is in 3NF if for all $X \rightarrow A$ in F^+, either (1) $A \in X$, or (2) X is a superkey of R, or (3) A is a member of some candidate key for R.

    (ii) On the other hand, R is in BCNF if (1) $A \in X$, or (2) X is a superkey for R.

    (iii) If a database schema is not well normalized, then even a conceptually simple update may require the locking of many objects, thus leading to poor performanance since the scheduler will have less opportunity to optimize transaction evaluation by interleaving tasks.