Distributed Information Systems: Spring Semester 2015

Quiz 2: Schema Fragmentation + Graph Databases

Student Name: _____ Date: 12 Mar 2015

Student ID: _____ Time: 11:15AM to 11:30AM

Total number of questions: 8

Each question has a single answer!

1. Which properties of a correct relational database fragmentation cannot be achieved simultaneously?

   ☒ *a*) Full disjointness and Reconstruction
   ☐ *b*) Completeness and Full disjointness
   ☐ *c*) Atomicity and Completeness
   ☐ *d*) Reconstruction and Completeness

2. Which of the following statements is **wrong**?

   ☐ *a*) The MinFrag algorithm's output is dependent on the order of its input.
   ☒ *b*) It is more effective to apply first a vertical fragmentation and then horizontal fragmentation, because a relation contains in general many more tuples than attributes.
   ☐ *c*) The semi-join operator applies a projection onto the attributes of one of the two input relations.
   ☐ *d*) Tuples that are accessed with the same frequency by all applications are grouped in the same horizontal fragment.

3. Which of the following statements is **true**?

   ☐ *a*) The result set size of the MinFrag algorithm cannot be smaller than the number of applications accessing the database.
   ☐ *b*) The MinFrag algorithm sorts the input set of simple predicates in order to optimize the execution cost.
   ☐ *c*) The result set size of the MinFrag algorithm is independent of the order the simple predicates are processed in the algorithm.
   ☐ *d*) In each iteration step of the MinFrag algorithm, a new simple predicate is either added to the result set or replaces at least one predicate from the current result set.

4. What can be a problem with horizontal fragmentation?

   ☒ *a*) Unbalanced workload when distributing the fragments
   ☐ *b*) Replication of data
   ☐ *c*) Fragmentation algorithm is NP-hard
   ☐ *d*) All of them can be a problem.

5. Which of the following is semi-structured data?

   ☐ *a*) RDF
   ☐ *b*) HTML
   ☐ *c*) XML with a schema
   ☒ *d*) All of them are semi-structured.

6. D is a data graph, and S is a schema graph for D. Which of the following is **wrong**?

   ☐ *a*) S simulates D.
   ☒ *b*) Each label in D is occurring in S.
   ☒ *c*) The number of nodes in S is smaller or equal to the number of nodes in D.
   ☐ *d*) There exists at least one simulation relationship.

7. Let $\{R_1 \dots R_n\}$ be the set of all simulation relationships between two graphs S_1 and S_2. Which property of simulation relationships identifies a uniquely defined simulation relationship R within that set:

☐ *a*) R is a rooted simulation
☐ *b*) R is a typed simulation
☒ *c*) R is a maximal simulation
☐ *d*) R is a deterministic simulation

8. Which statement about data guides is **wrong**?

☐ *a*) Every path occurs at most once.
☐ *b*) Different nodes in the data guide can contain the same data graph node.
☒ *c*) Different outgoing edges of a data guide node may have the same label.
☐ *d*) A data guide can contain cycles.