Global Questions

Q1: A validation methodologies allows researchers to back up their claims in an objective manner. It allows the reader of a paper to reproduce the results and conclusions of the author. This is a fundamental requirement for science and scientific research.

Q1a. List a few papers in your reading list and identify their validation methodologies.

Q1b. Do the paper describe their methodology in sufficient detail for ensuring reproducibility of their claims?

Q2: Which validation methodologies are commonly used in database literature to back up following kinds of claims:

1. correctness
2. completeness
3. worst case computational complexity
4. Pictograms can simplify conceptual models (e.g. ERD, UML)
5. Concurrency control for multi-diemnsional data: Predicate locks are not effective as multiple granularity for locks.
6. Buffer management: "Least Recently Used" is a reasonable buffer replacement policy.

Q3: Compare and Contrast the following pairs of concepts within simulation methodology. Hint: Define each term, identify the strengths and weakness of each.

1. Workload simulation: trace driven vs. Statistical Distribution driven
2. Performance measure: response time vs. throughput
3. Candidate comparison: statistical mean based vs. 95-th percentile based.

Compare and Contrast Papers

Q1: Compare and contrast query processing methods discussed in the following papers:

1. T1.2, J.M. Hellerstein and M.Stonebraker, Anatomy of a Database System

Q2: Compare and contrast components / component-groups used in DBMS architecture in the following papers:

1. T1.2, J.M. Hellerstein and M.Stonebraker, Anatomy of a Database System
2. UMN1, E.Lim, S.Hwang, J.Srivastava, d. Clements, M.Ganesh, Myriad: Design and Implementation of a Federated Database Prototype

Paper Specific Questions

(Q3.1) N. Beckmann, H. Kriegel, R. Schneider and B. Seeger, The R* Tree: An Efficient and Robust Access Method for Points and Rectangles.

Q1: Compare and contrast R-tree and R*-tree.

Q2: Consider following spatial dataset and alternative pagings. Which paging is preferred by R*-tree? Why?

Paging 1 – (R1, R2), (R3, R4, R5)
Paging 2 – (R2, R3), (R1, R4, R5)