NATIONAL INSTITUTE OF TECHNOLOGY, TIRUCHIRAPALLI - 620015

Advanced Database Management Systems – CS 606 **3 credits – Core -Theory**
M.Tech./CS **Session: II semester - Jan. 2018 – May 2018**

Instructor:

Dr. K. Viswanathan Iyer
Professor
Computer Science
CSE- 102 Ground Floor

e-mail: kvi[AT]nitt.edu
mobile: 9487487681

Pre-requisites:

- Knowledge of data structures, algorithms and programming language principles and computer organization is essential.
- A first level exposure to database syetsms will be assumed.

The following knowledge-base will be assumed though a rapid review on database related topics will be given with points to books and handouts:

Elementary ideas in computer organization, notion of data types, programming language constructs and features, good coding practices, order notions and time-complexity, elementary probability, combinatorics, familiarity with arrays, stacks, queues, linked lists.

The need for a computerised database system, high-level organization of an RDBMS, types of users and their roles, advantages of a database system, features and facilities in data definition and manipulation languages, relational algebra and calculus, some familiarity of SQL, client-server system architecture and associated technologies as well as business logic.

Course overview:

For organizing moderately large volumes of structured data (operational data arising for example, in airlines reservations, banking applications, human resource management) the relational database management technology (RDBMS) is still the preferred choice in the industry. This underscores the importance of this course when the Big Data revolution is the current buzzword. In the first about 5-8 hours the course will do a review of basic concepts in data structures and databases. The course then proceeds with a review of schema design for relational systems often conveniently ignored at a first level. The design aspects are covered only from a formal perspective so that the pitfalls in schema design can be avoided. RDBMS architecture is then reviewed at a high level. In doing so a quick overview of core SQL details is done through a series of graded examples. The course then moves to the foundations of query processing – this begins with relational algebra. Then algorithmic considerations to process selection, projection and join are outlined and the use of a various strategies and heuristics are examined. A quick overview of SQL will be provided in about 3/4 lectures. This is followed by the foundations in transaction processing. The course then continues with concurrency control techniques beginning with correctness of concurrent execution. Finally recovery techniques are covered and some recent trends in RDBMS are outlined.

Note: In the limited time no encyclopaedic coverage of the broad areas in databases will be attempted. Instead the references provided will be sufficient for practical purposes to gain a wider sprectrum of knowledge in the area as applicable to most industry needs. Also skill development via coding will not be emphasized during the course.

The following texts and other references/handouts will be used. Students should consult these references as well as other readable technical articles for more material and current trends on the topic Both these books give a good coverage of the planned syllabus.

1. R.Ramakrishnan and J.Gehrke, Database Management Systems, McGraw-Hill, 2003.

2. A.Silberschatz, H.F.Korth and S.Sudarshan, Database System Concepts, McGraw-Hill, 2006.

Students are also welcome to browse through the contents of the relevant syllabus book for unitwise listing of topics.

Course objectives and expected outcomes

1. The primary objective is to provide students with a second level introduction to database systems as seen from the perspectives of popular texts on the subject. A secondary objective is to provide insight into some commonly studied related emerging topics.

2. The expected outcome is that 45% of the students will be immediately ready to tackle challenges in the software industry with respect to the knowledge pertaining to this course. The next 45% of the students will find the course interesting enough to be motivated and get involved in the technical aspects covered in the course.

3. After the course about 60% of the students should be able to participate in formal database design. About 80% will be able to write SQL scripts independently.

4. About 60% of the students will be able to pick up associated technologies such as Java or Python, ODBC related packages, PHP.

5. A tertiary objective is to quickly introduce XML technology for integration with RDBMS.

6. More than 85% of the students will be ready to absorb, at a high level, any new market innovation related to transaction processing and database recovery.

Course teaching-learning activities:

A. Chalk-and-talk type lectures will be the main mode of content-delivery starting from 9 Jan. 2018. There will be three to four classes per week. As a policy, the announced time-table will be followed to make a total of 40 lectures. Lectures will be delivered with pointers to references. Sometimes handouts will be posted. Subject to availability of time-slots tutorial sesseions will be conducted on the prerequisites with no credit for attendance – this will be apart from the three hour per week schedule.

During the process of lectures students must refer to the suggested books (available in the library) and other references and handouts. The following topics will be dealt with in some details.

Preliminaries in defining a relational system. functional dependency and discovering them systematically, schema design via 3NF, BCNF, MVDs, hard problems in the design, a peek into modern approach to design.

Need for processing single queries optimally, SQL and equivalent relational algebra expressions, processing selection – file scan and other assumptions, processing projection, processing aggregate operations, processing join – use of hashing, equivalent transformations, heuristic query optimizations.

Characteristics of transactions, enhanced SQL provisions, modeling a transaction, need for concurrent executions, correctness of interleaved executions, cascadelessness and recoverability, locking basics, 2PL and its correctness, 2PL variants, performance of 2PL, notion of timestamps, basic timestamp ordering, optimistic concurrency control, dynamic databases.

System failures and software failures, recoverability techniques – UNDO, REDO protocol, some recent trends in database related technology.

B. For every one lecture an average student is expected to spend two hours. The supplementary readings will help in the understanding of the subject. During the first two weeks nonCS students are required to refer to suggested books to get the prerequisites.

Course assessment plan:

The assessment will consist of 1 quiz, 1 assignment, 2 tests, 1 end-of-semester examination.

The first quiz will also test the basics including the suggested prerequisites. The single assignment will be given in two parts, one part before the first test and one part before the second test. The weightages and durations are as under.

(a)	1 quiz (on pre-requisites):	10 marks	-	45 minutes
(b)	1 assignment in two/three parts:	10 marks	-	take about 3 hours time
(c)	2 tests:	20 + 20 = 40 marks	-	1 hour for each test
(d)	1 end-of-semester examination:	40 marks	-	3 hours

Note: (i) One compensatory assessment will be given – this will be for either a missed quiz or a test i.e., (a) or (c) and it will be of a better quality so that you are discouraged to be lethargic.

(ii) Late submissions of assignments will be suitably punished by award of lesser marks (first day delay - -2 marks; second day delay - -4 marks; third day delay: - -8 marks)

The exact dates will be announced in the first week in the class and will be posted. The first quiz will be conducted on 23 Jan. 2018 during the class time. Any changes will necessarily appear in the department notice board also.

The overall passing minimum will be computed as per senate directives. Letter grades as per institute policy will be assigned. The rules have provisions for any student to see the final examination papers.

Course feedback from students

Suggestions during the progress of the course are invited from all segments of the students. Constructive criticisms and comments are also welcome e.g., difficulties in prerequisites, tests, assignments, lectures. Students can leave the comments in the mailbox in the office.

Course policy – academic honesty, attendance, copying:

(a) Maintaining at least 75% attendance is recommended – institute attendance rules will apply in all cases. Only genuine reasons for attendance shortage of upto 25% will be allowed. Students with an attendance of 50% to 74% must attend mandatory classes before the final examination.

(b) Copying in assignments or tests will be viewed seriously. Students are expected to display a high degree of professionalism. Any unethical practice will attract punishment. Collaborating in assignments should be marked in the first page.

Sd/-
Dr.K.V.Iyer
Instructor

dtd. 9 Jan. 2018