UNIVERSITE LIBRE DE BRUXELLES
Faculté de Sciences Appliquées

January 2012

INFO-H-417 : Database System Architecture Course Information

The objectives in brief

In contrast to a typical introductory course in database systems where one learns to design and query relational databases, the goal of this course is to get a fundamental insight into the implementation aspects of database systems. In particular, we take a look under the hood of relational database management systems, with a focus on query and transaction processing.

With respect to query processing, we study the whole workflow of how a typical relational database management system optimizes and executes SQL queries. By having an in-depth understanding of the query-optimization-and-execution pipeline, one becomes more proficient in administering DBMSs, and hand-optimizing SQL queries for fast execution.

With respect to transaction processing we study how a typical relational database management systems ensures recovery from errors and controls concurrent access to the data.

Contacts

- Course responsible : Stijn Vansummeren (UB4.125, `stijn.vansummeren@ulb.ac.be`)
- Assistant : François Picalausa (UB4.133, `fpicalau@ulb.ac.be`)
- Course Web Page : `http://cs.ulb.ac.be/public/teaching/infoh417`

Schedule

Date	**Time**		**Content**	**Room**
Tue 14 Feb	14h–17h	Lect.1	Translation of SQL into the Relational Algebra	S.K.4.401
Tue 28 Feb	14h–17h	Ex. 1	Exercises on SQL-to-Relational-Algebra translation	S.K.4.401
		R.A. 1	Reading assignment : Physical Data Organization	
Tue 06 Mar	14h–17h	Ex. 2	Exercises on Logical Query Optimization	S.K.4.401
		Lect. 2	Logical Query Optimization	
Tue 13 Mar	14h–17h	Lect. 3	Index Structures (including introduction to hashing)	S.K.4.401
		Project	Project assignment + introduction	
Tue 20 Mar	14h–17h	Lect. 4	Multidimensional index structures &OLAP	S.K.4.401
Tue 27 Mar	14h–17h	Lect. 5	Physical operators	S.K.4.401
Tue 17 Apr	14h–17h	Ex 3.	Exercises on phyiscal operators	S.K.4.401
		Lect. 6	Cost-based plan selection	
Tue 24 Apr	14h–17h	Ex. 4	Exercises on cost-based plan selection	S.K.4.401
Tue 08 May	14h–17h	Ex. 5	Integrated exercises	S.K.4.401
		Demo 1	Demo : cost-based plan selection in practice	
		R.A. 2	Reading assignment : coping with system failures	
Tue 15 May	14h–17h	Ex. 6	Exercises on coping with system failures	S.K.4.401
		Lect. 7	Transaction processing : concurrency control	
Tue 22 May	14h–17h	Ex. 6	Exercises on concurrency control	S.K.4.401
		Demo 2	Demo 2 : Transaction processing in practice	
		Lect 8	More about transaction processing	
			Question & answer session	
Fri 1 June		**Deadline**	**Project deadline**	

Detailed Objectives

Upon successful completion of this course, the student masters the following competences :

1. Translating a given SQL expression into the Relational Algebra
2. Improving a relational algebra expression by, where possible, removing redundant joins in select-project-join subexpressions
3. Improving a relational algebra expression by, where possible, (a) replacing cartesian products by joins ; and (b) pushing selections and projections
4. Describing and being able to implement traditional secondary-memory index structures (BTrees, Hashing)
5. Being able to describe and demonstrate the shortcomings of traditional index structures with respect to multi-dimensional search keys. In addition, explaining the studied multi-dimensional indexes by means of an example
6. Describing the most important implementation algorithms (one-pass, sorting, hashing, index) for each of the relational algebra operators, as well as judging the cost of each operator, and knowing their limitations of applicability
7. Given a logical query plan and given base statistics about the size and distributions of the database relations, constructing a heuristically optimal physical query plan, by estimating the sizes of the intermediate results and correspondingly comparing the possible implementations. When joins can be reordered, choosing the order with the least cost.
8. Solving exercises on logging like the ones in sections 17.2.6, 17.3.5 en 17.4.4
9. Solving exercises on concurrency control like exercises 18.2.4, 18.2.5, 18.8.1, 18.8.2, en 18.9.1
10. Solving exercises on recoverability like exercises 19.1.1, 19.1.2 en 19.1.3
11. Being able to reconstruct the studied proofs

Method of evaluation

The project work contributes 6/20 points to the overall score, and the written exam contributes the remaining 14/20 points. Participation in both the project work and the written exam are mandatory requirements for passing the course.