

Plan de cours IFT723 – automne 2009

Sujets approfondis en bases de données

Enseignant	Date
Luc LAVOIE	2009-08-29
Statut	Version
En vigueur	1.1.0
Horaire	
Mardi, de 13:30 à 15:20 au D3-2032	
Vendredi, de 10:30 à 12h:20 au D3-2030	

1	Introduction	2
1.1	Objet et portée du document.....	2
1.2	Références	2
2	Présentation.....	3
2.1	Mise en contexte.....	3
2.2	Fiche signalétique	3
3	Contenu.....	4
4	Organisation.....	4
4.1	Modalités d'enseignement.....	4
4.2	Modalités d'évaluation.....	4
4.3	Calendrier	5
5	English synopsis	5

1 Introduction

1.1 Objet et portée du document

Le document décrit l'activité IFT723 « **Sujets approfondis en bases de données** » offerte au trimestre d'automne 2009. On y présente les objectifs, le contenu, l'organisation et les modalités d'évaluation du cours.

1.2 Références

[Date1997]

DATE, C. J. ; DARWEN, H.
A guide to the SQL standard.
4th ed., Addison-Wesley Inc., 1997.
ISBN 0-201-96426-0

[Date2007]

DATE, C. J. ; DARWEN, H.
Databases, types, and the relational model: The third manifesto.
3rd ed., Addison-Wesley Inc., 2008.
ISBN 0-321-39942-0

[Elmasri2007]

ELMASRI, Ramez ; NAVATHE, Shamkant B.
Fundamentals of database systems.
Fifth Edition, Pearson Addison Wesley, 2007.
ISBN 0-321-36957-2.

[Silberschatz2006]

SILBERSCHATZ, A. ; KORTH, H. ; SUDARSHAN, S.
Database Systems Concepts.
5th ed., McGraw-Hill, 2006, 1168 p.
ISBN 0072958863.

[Ullman2008]

ULLMAN, Jeffrey D. ; WIDOM, Jennifer ;
A First Course in Database Systems.
23rd ed., Prentice-Hall, 2008, 564 p.
ISBN 978-0-13-600637-4.

2 Présentation

2.1 Mise en contexte

Les bases de données jouent un rôle central dans le développement des systèmes informatiques. Elles permettent de stocker l'information relative à un domaine d'application, d'en préserver l'intégrité, de l'extraire en utilisant un langage de haut niveau, de traiter plusieurs transactions simultanément, de répartir les données, et d'assurer la sécurité et le recouvrement des données.

Le modèle relationnel domine depuis plus de 30 ans le domaine des bases de données, tant du point de vue théorique que du point de vue pratique. Il s'agit d'une des plus belles réussites de la recherche en informatique. On y retrouve l'élégance des mathématiques appliquées de manière efficace à un problème concret. Le modèle relationnel valut à son auteur, E. F. Codd, le *ACM Turing Award*. De par sa puissance, sa simplicité, son niveau d'abstraction, ses fondements mathématiques et son degré de pénétration du marché qui illustre son adéquation, le modèle relationnel constitue un cas de figure intéressant pour l'étude des bases de données.

Cette activité termine une série de quatre sur les bases de données (IFT187, IFT287, IFT487). On s'y consacre d'abord à une revue critique de la principale mise en oeuvre du modèle relationnel, SQL. Plusieurs améliorations sont ensuite analysées sur la base d'articles récents choisis au gré des participants et de l'enseignant.

2.2 Fiche signalétique

Objectif

Reconnaître les activités et les problèmes de la modélisation des données dans le contexte des bases de données; reconnaître les problèmes de recherche fondamentaux dans le domaine des bases de données.

Contenu

Analyse de différents modèles de données (réseau, relationnel, sémantique, etc.). Concepts fondamentaux : structures, contraintes, opérations. Conception des bases de données centralisées et distribuées. Étapes de la conception, modélisation conceptuelle, implantation, administration des bases de données (DBA). Répartition et allocation des données, concurrence, intégrité et recouvrement. Orientations futures : les machines BD (*database machines*), les systèmes de gestion des systèmes de bases de données intelligentes, les bases de données orientées objets telles que Object Store, O2 et Versant, ainsi que les bases de données déductives.

Préalable

IFT 487 ou l'équivalent.

3 Contenu

1. Introduction
2. Le modèle relationnel et le langage SQL
3. L'influence de l'orienté-objet
4. Analyse critique selon Date et Darwen
5. Tutorial D
6. Réponses aux critiques de Date et Darwen
7. Les problèmes découlant du modèle transactionnel
8. Streaming, Map-Reduce et modèle relationnel vertical
9. Le retour des modèles hiérarchiques et réseau
10. Choix de sujets contemporains

4 Organisation

4.1 Modalités d'enseignement

Les périodes de cours hebdomadaires serviront aux exposés théoriques et aux exemples. Les étudiants sont invités à participer directement aux cours par des présentations d'articles, de rapports de recherche et de rapports de synthèse.

4.2 Modalités d'évaluation

L'évaluation comporte deux contrôles, un travail de programmation, un essai et une présentation de 30 minutes. Les contrôles portent sur les sujets abordés par l'enseignant en cours. La présentation d'une durée approximative de 30 minutes est celle de l'essai préparé par l'étudiant.

Tableau 1 – Sommaire des évaluations

Évaluation	Valeur	Commentaire
Contrôle I	20 %	Individuel
Contrôle II	20 %	Individuel
Travail	20 %	Individuel
Essai	30 %	Individuel
Présentation	10 %	Individuel
Total	100 %	

Conformément aux articles 36, 37 et 38 du règlement facultaire d'évaluation des apprentissages¹, l'enseignant peut retourner à l'étudiante ou à l'étudiant tout travail non conforme aux exigences quant à la qualité de la langue et aux normes de présentation.

Toute situation de plagiat sera traitée en conformité, entre autres, avec l'article 8.1.2 du Règlement des études² de l'Université de Sherbrooke.

4.3 Calendrier

Tableau 2 – Calendrier des activités

N ^o	Semaine	Activité	Contenu	Lectures	Évaluations, remises, etc.
1	2008-08-31	cours	1, 2	Date2007	
2	2008-09-07	cours	3	articles	
3	2008-09-14	cours	4	Date2007	
4	2008-09-21	cours	4	Date2007	
5	2008-09-28	cours	4	Date2007	Énoncé du travail
6	2008-10-05	cours	5	Date2007	Choix du sujet de l'essai
7	2008-10-12	relâche	xx		Contrôle
8	2008-10-19	relâche	xx		
9	2008-10-26	relâche	xx		
10	2008-11-02	cours	6	Ullman2008, articles	Remise du travail
11	2008-11-09	cours	7	articles	Remise du plan de travail de l'essai
12	2008-11-16	cours	8	articles	
13	2008-11-23	cours	9	articles	Contrôle
14	2008-11-30	séminaire	10		
15	2008-12-07	relâche	xx		Remise de l'essai

Il y a deux séances hebdomadaires de cours, chacune d'une durée de deux heures. Le cours fait relâche durant trois semaines, du 12 octobre au 1^{er} novembre.

5 English synopsis

Presentation

Databases play a central role in the development of information systems. They store information relating to an application domain maintaining the integrity, handling multiple transactions simultaneously, distributing data, and ensuring security and data recovery.

¹ http://www.usherbrooke.ca/accueil/documents/politiques/pol_2500-008/pol_evaluation/sciences.html

² <http://www.usherbrooke.ca/programmes/etude>

The relational model dominates more than 30 years the field of databases, both theoretically and in practice. This is one of the greatest achievements in Information Science research. The elegance of mathematics is applied effectively to a concrete problem. The relational model earned its author, E. F. Codd, the ACM Turing Award. By its power, its simplicity, its level of abstraction, its mathematical foundations and its degree of market penetration that illustrates its appropriateness, the relational model is an interesting scenario for the study of databases.

This course completes a series of four databases courses (IFT187, IFT287, and IFT487). It is dedicated primarily to a critical review of the main implementation of the relational model, SQL. Several improvements are analyzed on the basis of recent papers selected at the discretion of the participants and the teacher.

Content

- 1. Introduction**
- 2. The relational model and the SQL language**
- 3. Object orientation influence**
- 4. Date and Darwen Analysis and Critics**
- 5. Tutorial D**
- 6. Some answers to Date and Darwen Critics**
- 7. Transactional model issues**
- 8. Streaming, Map-Reduce et vertical relational model**
- 9. Hierarchical et network models revival**
- 10. Electives topics**

Organisation

The periods are used for weekly presentations and examples. Students are invited to participate directly in the course by presentations of papers, research reports and summary reports.

There are two sessions per week, each for two hours.

The evaluation focuses on two controls, one homework and one report related to scientific papers to be agreed with the teacher. The report must also be presented as a 30 minutes conference.