



Département d'informatique
IGE 487 – Modélisation de bases de données
Plan d'activité pédagogique
Automne 2018

Enseignant :**Luc Lavoie**

Courriel : luc.lavoie@usherbrooke.ca
Local : D4-2006
Téléphone : 819-821-8000 (62015)
Site : <http://info.usherbrooke.ca/llavoie>
Disponibilité : sur rendez-vous

Responsables : Marc Frappier et Luc Lavoie

Horaire :

Lundi	08h30 à 10h20	Salle D4-2021
Mercredi	15h30 à 17h20	Salle D4-2021

Description officielle de l'activité pédagogique¹

Cibles de formation : Connaître les principes fondamentaux sous-jacents aux bases de données. Connaître les concepts de recherche d'information.

Contenu : Base de données orientée objets. Notations ODL, OQL, UML. Traduction d'un modèle orienté objet en un modèle relationnel. Modèle relationnel objet et relationnel étendu. Algèbre relationnelle. Transactions. Contrôle de la concurrence. Recouvrement. Optimisation des requêtes. Conception d'une base de données distribuée. Entrepôts et forage de données.

Crédits : 3

Organisation : 3 heures d'exposé magistral par semaine ;
0 heure d'exercice par semaine ;
6 heures de travail personnel par semaine.

Préalable : IFT 287

Particularités : Aucune

¹ <http://www.usherbrooke.ca/fiches-cours/ige487>

1 Présentation

Cette section présente les objectifs spécifiques et le contenu détaillé de l'activité pédagogique.

1.1 Mise en contexte

On a (trop) souvent tendance à oublier que l'information est l'objet² d'étude de l'informatique. Une des questions centrales est celle de la représentation de l'information sous la forme de données et du maintien de son intégrité en cours de traitement. Se greffent à ces considérations fondamentales plusieurs problèmes liés aux technologies utilisées (concurrence d'accès, par exemple) et au contexte d'exploitation (contrôle d'accès, par exemple).

Depuis la fin de la Deuxième Guerre mondiale, des développements technologiques toujours plus nombreux et plus rapides ont amené les spéculations scientifiques (Turing) et sociales (Orwell) au rang des phénomènes concrets. Il en découle que les questions scientifiques, sociales, technologiques et commerciales liées à la gestion de l'information et à celle des données (et donc aux bases et entrepôts de données) sont au centre du développement et de l'évolution des systèmes logiciels devenus essentiels à la bonne gestion des organisations et au centre de la mission de plusieurs d'entre elles. La gestion de l'information est aussi devenue essentielle à l'avancement de plusieurs domaines de connaissances dont l'écologie, la génomique, la climatologie, la démographie et l'économique.

À partir des bases établies grâce aux activités IFT 187 et IFT 287, l'activité IGE 487 s'intéresse donc à :

- la qualité de la représentation de l'information et donc à son intégrité (théorie relationnelle) ;
- la dimension temporelle des données (logique temporelle par intervalle) ;
- l'accès aux données (systèmes transactionnels et propriétés ACID) ;
- l'interprétation des données (systèmes analytiques : entrepôts, étoiles, flocons et cubes) ;
- l'exploitation des données (journalisation, sauvegarde, répartition, distribution).

D'autres aspects pourront être traités tels que : la théorie corelationnelle (l'origine des produits dits noSQL), le contrôle d'accès (sécurité) et la modélisation ontologique.

1.2 Objectifs spécifiques

À la fin de l'activité, l'étudiant doit être capable :

1. d'évaluer et de prendre en compte l'impact des choix technologiques gouvernant la mise en oeuvre des systèmes de gestion de bases de données relationnelles ;
2. de mettre en oeuvre et de faire évoluer une base de données transactionnelle de moyenne envergure à l'aide d'outils contemporains ;
3. de concevoir et mettre en oeuvre une base de données analytique selon les règles de l'art (étoiles et flocons) ;
4. d'appliquer les principaux modèles de temporalité des problèmes simples.

² Parmi les nombreuses causes, il convient de mentionner l'absence de mot pour désigner l'informatique en anglais. Si un germanophone, un hispanophone ou un russophone ne peut « oublier » la centralité du concept d'information puisqu'il se trouve à la racine du nom de la « Science du traitement automatique et rationnel de l'information », l'anglophone peut être porté à l'oublier lorsqu'il ne perçoit que la technologie derrière la « Computer science ». Comme si la physique était la science du voltmètre, la chimie celle du spectrographe de masse et la biologie celle du microscope.

1.3 Contenu détaillé

Thème	Contenu	Heures ¹	Objectifs	Travaux ²
1.	Revue et approfondissement de la théorie relationnelle (propriétés des relations, algèbre et calcul relationnels, interchangeabilité des relations et des vues)	4	1	X
2.	Revue et approfondissement de la conception relationnelle (formes normales 1FN à 6FN et FNBC, construction direction du schéma logique ou par l'intermédiaire de modèles conceptuels)	4	1	X
3.	Optimisation des requêtes et application (arbres d'expression, permutations heuristiques, réduction, extension, fonctions de cout ; utilisation des plans d'exécution en SQL)	4	2	X
4.	Mise en oeuvre des propriétés ACID (modèle transactionnel : journal, granularité, types d'équivalence, caractérisation des problèmes, récupération ; gestion de la concurrence : algorithmes de verrouillage et d'estampillage)	8	2	
5.	Journalisation externe et sauvegarde (spécification des exigences, survol des principales méthodes)	2	2	
6.	Répartition et distribution (spécification des exigences, survol des principales méthodes)	2	2	
7.	Revue et approfondissement des schémas transactionnels (isolation des relations de base et développement de schémas virtuels ; spécialisation des schémas en fonction de l'abstraction et du contrôle d'accès ; développement des fonctions de modification abstraites)	4		
8.	Modèles d'entrepôts partiellement temporalisés (étoiles, flocons et cubes ; entrepôts dimensionnels : faits et dimensions, construction des clés et intégration partielle de la temporalité, opérations spécifiques [drill down, drill up, drill through, drill across] ; problématiques des données manquantes ; problématique de la modification ; contrôle de la redondance et de l'intégrité ; problématique de l'alimentation [ETL vs ELT] ; problématique de l'évolution)	8	3	X
9.	Temporalité et intervalles (ITL) (point, intervalle, opérateurs de Allen ; modèle discret du temps : instant, période, opérateurs pack et unpack)	4	4	X
10.	Modèles d'entrepôts bitemporalisés (BCDM, DDLM) (introduction au modèle BCDM de Snodgrass ; introduction au modèle DDLM de Date, Darwen et Lorentzos)	8	3, 4	X
11.	Sujets électifs (par exemple : modélisation ontologique, index temporels)	4	1-4	
Total		52		

1. Répartition des heures combinées de cours, d'exercices et de travaux dirigés sur la base de 13 semaines à raison de 4 heures par semaine, pour un total de 52 heures.
2. Les travaux doivent couvrir minimalement les sujets marqués d'un X dans la colonne « Travaux » du tableau précédent. Il peut s'agir d'au moins quatre travaux pratiques ou un projet comprenant au moins quatre jalons intégrant l'ensemble des apprentissages.

2 Organisation

Cette section présente la méthode pédagogique, le calendrier, le barème et la procédure d'évaluation ainsi que l'échéancier des travaux.

2.1 Méthode pédagogique

Des exposés magistraux présentent et motivent chacun des thèmes, tant sur le plan des connaissances théoriques que de la méthode de mise en oeuvre. Des lectures autonomes complètent ces présentations et sont essentielles à l'atteinte des objectifs. Un projet, comprenant au moins 4 jalons, permet d'intégrer les apprentissages.

2.2 Calendrier

Le calendrier est structuré selon une répartition approximative des heures de cours à raison de quatre heures par semaine. Au besoin, cette modulation peut être adaptée en fonction des besoins d'approfondissement exprimés par les étudiants.

Tableau 1 – Planification des activités et des lectures

N°	Semaine	Activités	Contenu	Travaux
1.	2017-08-27	cours	1	
2.	2017-09-03	cours	2	
3.	2017-09-10	cours	7	J1
4.	2017-09-17	cours	3	
5.	2017-09-24	cours	4	J2
6.	2017-10-01	cours	4	
7.	2017-10-08	début des examens le 6		examen 1
8.	2017-10-15	relâche universitaire		
9.	2017-10-22	cours	5, 6	
10.	2017-10-29	cours	8	
11.	2017-11-05	cours	8	J3
12.	2017-11-12	cours	9	
13.	2017-11-19	cours	10	
14.	2017-11-26	cours	10	J4
15.	2017-12-03	cours	11	
16.	2017-12-10	dernier cours le 10 ; début des examens le 11 ; fin des examens le 21	--	projet examen 2

Le premier cours est programmé le lundi 27 août et le dernier le lundi 10 décembre. Les dates de jalons du projet sont sujettes à changement en fonction du rythme du cours. Il n'y a pas de cours le lundi 3 septembre (congé universitaire) ni le lundi 1^{er} octobre (élections). Les dates d'examen seront fixées ultérieurement par la Faculté des sciences.

2.3 Évaluation

En plus des deux examens individuels, l'évaluation porte sur un projet réalisé en équipe de deux à trois personnes. Le projet est divisé en quatre livraisons préliminaires correspondant aux jalons J1 à J4 et une livraison finale à remettre le jour de l'examen final.

Les examens sont placés sous la responsabilité de la Faculté des sciences et organisés par elle. La durée des examens est de trois heures. La documentation personnelle est limitée à une seule feuille recto verso au format lettre ou A4 ; l'usage d'appareils informatiques, électroniques ou de communication (ordinateur, calculatrice, téléphone, etc.) est interdit sauf si l'examen a lieu dans un laboratoire auquel cas seuls les équipements du laboratoire peuvent être utilisés selon des conditions qui seront alors précisées.

Tableau 2 – Sommaire des évaluations

Évaluation	Valeur	Commentaire
examen 1	30 %	individuel
examen 2	40 %	individuel et récapitulatif
travaux	30 %	en équipe de trois à quatre personnes
Total	100 %	

Tout étudiant, toute étudiante, qui omet de remettre un travail au moment prescrit doit rencontrer l'enseignant afin de déterminer une nouvelle date de remise. Une pénalité de 10 % par jour de retard est imposée.

L'évaluation est faite en tenant compte de la clarté des documents et du respect de la méthodologie du génie logiciel. Conformément au Règlement facultaire d'évaluation des apprentissages³, 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 avec le Règlement des études⁴ de l'Université de Sherbrooke.

En cas de circonstances extraordinaires au-delà du contrôle de l'Université de Sherbrooke et sur décision de celle-ci, l'évaluation des apprentissages de cette activité est sujette à changement.

2.4 Échéancier des travaux

Voir le calendrier (section 2.2).

2.5 Utilisation d'appareils électroniques et du courriel

Dans la présente activité, l'usage de téléphones cellulaires, de tablettes ou d'ordinateurs est autorisé. Cette permission peut être retirée en tout temps si leur usage perturbe le déroulement des prestations ou entraîne des abus. Toute utilisation d'appareils de captation de la voix ou de l'image nécessite la permission de l'enseignant qui en fixera les conditions.

Note : L'utilisation du courriel est recommandée pour prendre rendez-vous avec l'enseignant ou lui adresser des questions.

3 Matériel nécessaire pour l'activité

Le plan de l'activité et les présentations utilisées en cours sont disponibles sur les sites de l'enseignant⁵. Les manuels de base recommandés sont [Adamson 2011] [Date 2014] et [Elmasri 2016].

4 Références

Cette section présente les principales références utilisées par l'enseignant ou dont il recommande la lecture en tout ou en partie.

4.1 Références essentielles

[Elmasri 2016]

Ramez ELMASRI, Shamkant B. NAVATHE ;

Fundamentals of database systems.

7th edition, Pearson, 2016. ISBN 978-0-13-397077-7.

³ https://www.usherbrooke.ca/sciences/fileadmin/sites/sciences/Etudiants_actuels/Informations_academiques_et_reglements/2017-10-27_Reglement_facultaire_-_evaluation_des_apprentissages.pdf

⁴ <https://www.usherbrooke.ca/registraire/droits-et-responsabilites/reglement-des-etudes/>

⁵ <http://info.usherbrooke.ca/llavoie/enseignement/IGE487>
<http://www.genilog.org/llavoie/enseignement/IGE487>

[IGE 487]

Luc LAVOIE ;
IGE 487 – Modélisation de bases de données.
Faculté des sciences, Université de Sherbrooke, septembre 2018 ;
<http://info.usherbrooke.ca/lavoie/enseignement/IGE487>

4.2 Références utiles

[Adamson 2010]

Christopher ADAMSON;
Star Schema - The Complete Reference.
McGraw-Hill, 2010. ISBN 978-0-12-800631-3.

[BFO 2015]

Robert ARP, Barry SMITH, Andrew D. SPEAR;
Building ontologies with Basic Formal Ontologies.
The MIT Press. Cambridge (MA) USA. 2015.

[Date 2011]

Chris J. DATE;
SQL and Relational Theory: How to Write Accurate SQL Code.
O'Reilly, 2011. ISBN 978-1-4493-1640-2.

[Date 2012]

Chris J. DATE;
Database Design and Relational Theory: Normal Forms and All That Jazz.
O'Reilly, 2012. ISBN 978-1-4493-2801-6.

[Date 2014]

Chris J. DATE, Hugh DARWEN, Nikos A. LORENTZOS;
Time and Relational Theory: Temporal Databases in the Relational Model and SQL.
Morgan Kaufmann (Elsevier), 2014. ISBN 978-1-449-31640-2.

[Jiang 2015]

Jiang, Bin;
Constructing Data Warehouses with Metadata-Driven Generic Operators, and More Architecture, Methodology, and Paradigm, Concepts, Algorithms, and Operators, Principles, Recommendations, and Exercises.
2nd edition, DBJ Publishing, 2015. ISBN 978-15086873-13.

[Johnston 2014]

Tom JOHNSTON, Randall WEIS;
Managing Time in Relational Databases: How to Design, Update and Query Temporal Data.
Morgan Kaufmann, 2010. ISBN 978-0123750419.

[Snodgrass 1999]

Richard T. SNODGRASS;
Developping time-oriented database applications in SQL.
Morgan Kaufmann, 1999. ISBN 978-1558604360
(épuisé, mais disponible sur <http://www.cs.arizona.edu/people/rts/tdbbook.pdf> en date du 2015-08-11).

[Ullman 2008]

Jeffrey D. ULLMAN, Jennifer WIDOM;
A First Course in Database Systems.
3rd edition, Prentice-Hall, 2008. ISBN 978-0-13-600637-4.

4.3 Références aux SGBD utilisés pour les exemples, exercices et travaux

MariaDB (en anglais, 2018-08-218)

** <https://mariadb.com/kb/en/library/documentation/>

Oracle (en anglais, 2018-08-18)

** https://docs.oracle.com/cd/E11882_01/index.htm

PostgreSQL (en anglais, 2018-08-18)

** <https://www.postgresql.org/docs/10/static/index.html>

PostgreSQL (*en français*, 2018-08-18)

** <https://docs.postgresqlfr.org>



L'intégrité intellectuelle passe, notamment, par la reconnaissance des sources utilisées. À l'Université de Sherbrooke, on y veille!

Extrait du Règlement des études (Règlement 2575-009)

9.4.1 DÉLITS RELATIFS AUX ÉTUDES

Un délit relatif aux études désigne tout acte trompeur ou toute tentative de commettre un tel acte, quant au rendement scolaire ou une exigence relative à une activité pédagogique, à un programme ou à un parcours libre.

Sont notamment considérés comme un délit relatif aux études les faits suivants :

- a) commettre un plagiat, soit faire passer ou tenter de faire passer pour sien, dans une production évaluée, le travail d'une autre personne ou des passages ou des idées tirés de l'œuvre d'autrui (ce qui inclut notamment le fait de ne pas indiquer la source d'une production, d'un passage ou d'une idée tirée de l'œuvre d'autrui);
- b) commettre un autoplgiat, soit soumettre, sans autorisation préalable, une même production, en tout ou en partie, à plus d'une activité pédagogique ou dans une même activité pédagogique (notamment en cas de reprise);
- c) usurper l'identité d'une autre personne ou procéder à une substitution de personne lors d'une production évaluée ou de toute autre prestation obligatoire;
- d) fournir ou obtenir toute aide non autorisée, qu'elle soit collective ou individuelle, pour une production faisant l'objet d'une évaluation;
- e) obtenir par vol ou toute autre manœuvre frauduleuse, posséder ou utiliser du matériel de toute forme (incluant le numérique) non autorisé avant ou pendant une production faisant l'objet d'une évaluation;
- f) copier, contrefaire ou falsifier un document pour l'évaluation d'une activité pédagogique;

[...]

Par plagiat, on entend notamment :

- Copier intégralement une phrase ou un passage d'un livre, d'un article de journal ou de revue, d'une page Web ou de tout autre document en omettant d'en mentionner la source ou de le mettre entre guillemets;
- reproduire des présentations, des dessins, des photographies, des graphiques, des données... sans en préciser la provenance et, dans certains cas, sans en avoir obtenu la permission de reproduire;
- utiliser, en tout ou en partie, du matériel sonore, graphique ou visuel, des pages Internet, du code de programme informatique ou des éléments de logiciel, des données ou résultats d'expérimentation ou toute autre information en provenance d'autrui en le faisant passer pour sien ou sans en citer les sources;
- résumer ou paraphraser l'idée d'un auteur sans en indiquer la source;
- traduire en partie ou en totalité un texte en omettant d'en mentionner la source ou de le mettre entre guillemets ;
- utiliser le travail d'un autre et le présenter comme sien (et ce, même si cette personne a donné son accord);
- acheter un travail sur le Web ou ailleurs et le faire passer pour sien;
- utiliser sans autorisation le même travail pour deux activités différentes (autoplgiat).

Autrement dit : mentionnez vos sources
