



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
DEPARTAMENTO DE CIENCIA DE LA COMPUTACIÓN

IIC3413 — Implementación de Sistemas de Base de Datos

Programa de Curso

Lunes 9 de Marzo de 2020

Profesor: Cristian Riveros (cristian.riveros@uc.cl, oficina 3-S, DCC).
Ayudantes: Martín Muñoz (mmunos@uc.cl).
 Andrés Quintana (afquintana@uc.cl).
Clases: Lunes y miércoles módulo 3, sala K301.
Ayudantías: Viernes módulo 3, sala K301.
Atención de alumnos: A toda hora, preferentemente después de clases.
Sitio Web: Canvas / Implementación de Sistemas de Base de Datos.

Descripción

Un sistema de base de datos está en el corazón de toda aplicación computacional moderna. Estos sistemas son los actores principales encargados de extraer, manejar y actualizar grandes volúmenes de datos eficientemente. Si bien las estructuras de datos y algoritmos pueden variar de un sistema a otro, sus principios han alcanzado un gran nivel de madurez. Varias de sus componentes como, por ejemplo, el almacenamiento de los datos, la optimización de consultas, o la administración de transacciones, son parte fundamental de su arquitectura y son muchas veces transversales a todos los sistemas comerciales.

Durante este curso, los alumnos estudiarán los principios de la construcción de un sistema de base de datos. Esto incluye entender las problemáticas y soluciones relacionadas con el manejo de grandes cantidades de datos. Estas técnicas comprenden las usadas en sistemas comerciales modernos, como así también las últimas propuestas en investigación.

Objetivos

Al finalizar el curso el alumno será capaz de:

1. Conocer el funcionamiento interno de un sistema de base de datos.
2. Aplicar distintas estructuras de datos y algoritmos para procesar grandes volúmenes de datos.
3. Diseñar e implementar las distintas componentes de un sistema de base de datos moderno.
4. Evaluar la eficiencia de estructuras de datos y algoritmos aplicados a problemas en base de datos.

Contenidos

Repaso de un sistema de bases de datos

1. Modelo relacional, consultas SQL, álgebra relacional.

2. Arquitectura de un sistema de bases de datos relacional, proceso de optimización de consultas SQL.

Almacenamiento e índices

1. Almacenamiento de datos: dispositivos de almacenamiento secundario, jerarquía de almacenamiento, tiempo de acceso a disco, modelo de computación Input/Output, acceso optimizado a disco.
2. Sistema de archivos: planificación y acceso a páginas, manejo de buffers, organización de archivos.
3. Estructura de índices: índices basados en árboles, basados en hashing, multi-dimensionales.

Evaluación y optimización de consultas

1. Evaluación de operadores: ordenamiento externo de datos, algoritmo de múltiples pasadas, ordenamiento basado en índices, selección, proyección, joins, operaciones de conjuntos, agregación, impacto del buffer, pipelining, blocking.
2. Optimización de consultas: sistema de catálogos, estimación del tamaño de los resultados,

histogramas, equivalencias de operadores relacionales, planes de consulta, estimación de costos, consultas anidadas, algoritmos de optimización de joins.

Consultas conjuntivas

1. Definición, homomorfismo. NP-completitud.
2. Consultas acíclicas, caracterizaciones, algoritmo de Yannakakis, enumeración constant-delay.
3. Tree-width, hypertree-width, caracterizaciones.

Optimalidad en el peor caso

1. Cota AGM, demostración.
2. Consultas multijoin, algoritmos eficientes en el peor caso,
3. Algoritmos para multi-intersección, Leapfrog-trie join.
4. Representación compacta de consultas.

Evaluación

La evaluación se realizará en base a laboratorios, controles y un examen final.

Laboratorios. Los laboratorios consisten en ejercicios individuales donde los alumnos deben implementar y evaluar algoritmos y estructuras de datos de una base de datos de grafos. Cada laboratorio deberá ser solucionado individualmente por cada alumno. Los enunciados y entrega de los laboratorios serán según el siguiente calendario:

	Publicación enunciado	Entrega
Laboratorio 1	Viernes 20 de Marzo	Viernes 3 de Abril
Laboratorio 2	Miércoles 15 de Abril	Miércoles 29 de Abril
Laboratorio 3	Viernes 1 de Mayo	Viernes 15 de Mayo
Laboratorio 4	Viernes 22 de Mayo	Viernes 5 de Junio
Laboratorio 5 (Opt)	Viernes 12 de Junio	Viernes 26 de Junio

Cada laboratorio tendrá un plazo de entrega de dos semanas luego de la publicación del enunciado. Como regla general, cada viernes después de la publicación de un enunciado se realizará una ayudantía donde se entregarán los detalles necesarios para completar cada laboratorio. La solución de cada laboratorio deberá ser enviada a mas tardar a las 23:59 horas del día de entrega. La entrega del laboratorio es obligatoria. El único laboratorio opcional es el laboratorio 5, el cual puede reemplazar la menor nota entre los laboratorios 1 al 4.

Controles y Examen. Durante el curso se realizarán tres controles y un examen en las siguientes fechas:

	Fecha
Control 1	Miércoles 15 de Abril
Control 2	Miércoles 13 de Mayo
Control 3	Miércoles 10 de Junio
Examen	Jueves 9 de Julio

Los controles son optativos y el examen es obligatorio. Los controles se realizarán durante las horas de clases y tendrán la duración de todo el módulo.

Aprobación del curso. El promedio de los laboratorios (**PL**) se calculará como el promedio aritmético entre todos los laboratorios. El promedio **PE** de los tres controles (C_1, C_2, C_3) y examen (E) se calculará según la siguiente formula

$$\mathbf{PE} = \max \left\{ E, \frac{C_1 + C_2 + C_3 + E - \min\{C_1, C_2, C_3\}}{3} \right\}$$

La nota final (**NF**) se calculará como:

$$\mathbf{NF} = 0,5 \cdot \mathbf{PL} + 0,5 \cdot \mathbf{PE}$$

El curso se aprueba si, y solo si, todas las siguiente condiciones se cumplen:

- promedio de laboratorios mayor o igual a 4,0 (**PL** \geq 4,0),
- promedio de controles y examen mayor o igual a 4,0 (**PE** \geq 4,0) y
- nota final mayor o igual a 4,0 (**NF** \geq 4,0).

En caso de no aprobar, la nota final del curso se calculará como $\min\{\mathbf{NF}, 3,9\}$.

Política de copia. En caso de copia (ver documento adjunto “Código de Honor de la Escuela de Ingeniería”) se tomarán las medidas dispuestas por la política de integridad académica del Departamento de Ciencia de la Computación (ver documento adjunto “Política de Integridad Académica del Departamento de Ciencia de la Computación”).

Comunicación digital

La vía de comunicación sobre noticias, material de clases y notas será:

Canvas / Implementación de Sistemas de Base de Datos.

Para preguntas del curso sobre contenidos o laboratorios se habilitará el foro del Canvas. En caso de tener preguntas personales pueden escribir a:

iic3413@ing.puc.cl

Este correo (alias) esta dirigido al profesor y ayudantes del curso y se buscará responder las dudas lo antes posible. Preferentemente, se sugiere enviar todas las dudas al foro del curso. En caso de tener preguntas por problemas personales relacionados al curso, escribir directamente al correo del profesor.

Bibliografía

Durante el curso se recomiendan los siguientes libros de estudio:

- *Database Systems: The Complete Book*. Hector Garcia-Molina, Jeffrey Ullman y Jennifer Widom. Prentice Hall, 2009 (2nd ed).
- *Database Management Systems*. Raghu Ramakrishnan y Johannes Gehrke. McGraw-Hill, 2003 (3rd ed).
- *Database System Concepts*. Abraham Silberschatz, Henry Korth y S. Sudarshan. McGraw-Hill, 2010 (6th ed).
- *Database Design and Implementation*. Edward Sciore. Wiley, 2008 (1er ed).

COMPROMISO DE CODIGO DE HONOR

Este curso adscribe el Código de Honor establecido por la Universidad, el que es vinculante. Todo trabajo evaluado en este curso debe ser propio. En caso de que exista colaboración permitida con otros alumnos, el trabajo deberá referenciar y atribuir correctamente dicha contribución a quien corresponda. Como alumno es su deber conocer el Código de Honor (www.uc.cl/codigodehonor).

Política de Integridad Académica del Departamento de Ciencia de la Computación

Se espera los alumnos de la Escuela de Ingeniería de la Pontificia Universidad Católica de Chile mantengan altos estándares de honestidad académica, acorde al Código de Honor de la Universidad. Cualquier acto deshonesto o fraude académico está prohibido; los alumnos que incurran en este tipo de acciones se exponen a un Procedimiento Sumario. Es responsabilidad de cada alumno conocer y respetar el documento sobre Integridad Académica publicado por la Dirección de Pregrado de la Escuela de Ingeniería (Disponible en SIDING, en la sección Pregrado/Asuntos Estudiantiles/Reglamentos/Reglamentos en Ingeniería/Integridad Académica).

Específicamente, para los cursos del Departamento de Ciencia de la Computación, rige obligatoriamente la siguiente política de integridad académica. Todo trabajo presentado por un alumno para los efectos de la evaluación de un curso debe ser hecho **individualmente** por el alumno, **sin apoyo en material de terceros**. Por “trabajo” se entiende en general las interrogaciones escritas, las tareas de programación u otras, los trabajos de laboratorio, los proyectos, el examen, entre otros.

En particular, si un alumno copia un trabajo, o si a un alumno se le prueba que compró o intentó comprar un trabajo, **obtendrá nota final 1.1 en el curso** y se solicitará a la Dirección de Pregrado de la Escuela de Ingeniería que no le permita retirar el curso de la carga académica semestral. Por “copia” se entiende incluir en el trabajo presentado como propio, partes hechas por otra persona. En caso que corresponda a “copia” a otros alumnos, la sanción anterior se aplicará a todos los involucrados. En todos los casos, se informará a la Dirección de Pregrado de la Escuela de Ingeniería para que tome sanciones adicionales si lo estima conveniente.

Obviamente, está permitido usar material disponible públicamente, por ejemplo, libros o contenidos tomados de Internet, **siempre y cuando se incluya la referencia correspondiente**.

Lo anterior se entiende como complemento al Reglamento del Alumno de la Pontificia Universidad Católica de Chile:

<http://admisionyregistros.uc.cl/alumnos/informacion-academica/reglamentos-estudiantiles>

Por ello, es posible pedir a la Universidad la aplicación de sanciones adicionales especificadas en dicho reglamento.