

Programa de Curso

Agosto de 2018

Profesor:	Adrián Soto assoto@uc.cl
Ayudantes	Isidora Vizcaya, Thomas Reisenegger, Martín Álamos, Romano Fenzo, Antonio Ossa Fernando Pieressa, Stephanie Chau, Luis Felipe Silva, Wenyi He Yeng, Kevin Johnson
Clases:	lunes y miércoles, módulo 5. Sala A7
Horario de Ayudantías:	viernes, módulo 5. Sala K204

Descripción

Las bases de datos forman parte del núcleo del desarrollo de aplicaciones comerciales modernas, y son indispensables para cualquier aplicación que requiera almacenar grandes volúmenes de datos, actualizarlos y consultarlos de manera eficiente. El propósito de este curso es introducir al alumno en el diseño y uso de los sistemas de bases de datos, específicamente bases de datos relacionales.

Objetivo General

Durante el curso el alumno aprenderá la teoría detrás del modelo relacional, incluyendo el lenguaje SQL (hoy estándar para consultar bases de datos relacionales). Sabrá aplicar esta teoría al diseño y uso de bases de datos relacionales. Además, conocerá como interpretar y consultar bases de datos noSQL. Al final del curso el alumno podrá diseñar y manejar una base de datos relacional en un ambiente real. También conocerá los modelos teóricos usados en el estudio de las bases de datos, podrá analizar y comparar distintos esquemas y consultas de acuerdo a su utilidad práctica, e interpretar las ventajas y desventajas de los modelos de datos relacionales en comparación con modelos noSQL.

Competencias e indicadores de desempeño

Competencia	Indicadores de Desempeño
1. Entender qué es un sistema de bases de datos, y conocer las ventajas y desventajas de los modelos de datos relacional y noSQL.	<ol style="list-style-type: none">1. El alumno conoce la necesidad de implementar y usar sistemas de bases de datos.2. Entiende la diferencia entre modelos de datos, esquema de datos e instancia de un esquema.3. Distingue entre distintos modelos de datos, conoce que hay distintas alternativas, sus ventajas y desventajas.

2. Diseñar bases de datos relacionales para aplicaciones basadas en casos reales	<ol style="list-style-type: none"> 1. Es capaz de crear un esquema de datos a partir de una descripción verbal. 2. Abstrae y optimiza requerimientos mediante un modelo ER. 3. Conoce como generar un esquema relacional a través de un modelo ER. 4. Compara distintos esquemas de acuerdo a los requerimientos. 5. Aplica la teoría de formas normales a la comparación y/o optimización de esquemas.
3. Usar el lenguaje de consulta SQL para bases de datos relacionales	<ol style="list-style-type: none"> 1. Conoce la sintaxis y la semántica de SQL. 2. Escribe consultas simples. 3. Establece un plan lógico frente a una consulta, y lo expresa en SQL. 4. Puede comprender el significado de una consulta compleja y establecer su plan lógico.
4. Comprender conceptos esenciales en teoría de bases de datos, especialmente el Álgebra Relacional (AR), y conoce como se usa para estudiar las limitaciones de las bases de datos	<ol style="list-style-type: none"> 1. Conoce la sintaxis y la semántica del Álgebra Relacional. 2. Puede transformar consultas de SQL a AR, y vice versa. 3. Entiende el concepto consultas no expresables en SQL o AR, y puede dar ejemplos. 4. Conoce el concepto de plan de consultas, sabe el papel del AR en la optimización de consultas.
5. Conocer nociones básicas del paradigma noSQL, búsqueda de texto y herramientas de análisis de datos	<ol style="list-style-type: none"> 1. Conoce el paradigma key-value para almacenar datos y el sistema MongoDB. 2. Conoce los desafíos detrás de la búsqueda en textos. 3. Realiza consulta simples a la web semántica. 4. Conoce herramientas clásicas de análisis de datos.
6. Comprender los desafíos y problemas más importantes de la industria de bases de datos actuales. Discutir y saber plantear soluciones de diseño e implementación de bases de datos.	<ol style="list-style-type: none"> 1. Comprende los conceptos detrás de la arquitectura de un sistema de bases de datos. 2. Conoce y usa conceptos avanzados del lenguaje SQL, como vistas, transacciones, procedimientos almacenados, cursores y serialización. 3. Implementa una plataforma Web para acceder a una base de datos, y mostrar el resultado de las consultas. 4. Diseña soluciones para un caso real del problema de integración de datos. 5. Entiende cómo un buen sistema de bases de datos puede ayudar a tomar decisiones estratégicas a nivel industrial.

Contenidos

Por qué un sistema de bases de datos

1. Importancia de las bases de datos.
2. Sistemas de bases de datos versus almacenamiento ad-hoc.
3. Arquitectura básica de los sistemas de bases de datos.
4. Modelos de datos: XML, key-value, relacional.

Introducción al modelo relacional

7. Conceptos básicos.
8. SQL básico.

Modelamiento (diseño)

9. Creación de Tablas.
10. Lenguajes de modelamiento; diagramas E/R.
11. Principios de diseño.
12. Desde los diagramas E/R a esquemas relacionales.
13. Modelado de restricciones, formas normales.

SQL avanzado

14. Outerjoins y valores nulos.

15. Vistas.

16. Procedimientos almacenados y cursores.

17. Transacciones y control de concurrencia.

18. Herramientas de gestión.

Conceptos teóricos

20. El Álgebra relacional.

21. Limitaciones de SQL: valores nulos, clausura transitiva, etc.

Indexación y optimización

22. Procesamiento de operadores relacionales.

23. Optimización de consultas relacionales.

24. Índices en memoria secundaria.

25. Índices invertidos y búsqueda en texto

Otros modelos de datos

25. Paradigma key-value

26. Análisis de datos con Pandas

27. Web Semántica

Metodología

Clases expositivas y de discusión activa, ocasionalmente entre pares. Un proyecto semestral de desarrollo de una base de datos cuya información se revelará a medida que avanza el curso.

Evaluación

La evaluación se realizará en base a:

- 3 Interrogaciones, cuyo objetivo es evaluar conocimientos teóricos.
- Examen, con el mismo objetivo de evaluar conocimientos teóricos.
- Un proyecto semestral con 4 evaluaciones parciales. El objetivo del proyecto es que el alumno desarrolle una base de datos para una aplicación basada en un caso real. Las etapas avanzadas del proyecto introducirán al alumno otras problemáticas reales de las bases de datos.

La **nota de interrogaciones y exámenes (NIE)** corresponde al promedio de las interrogaciones y el examen, en donde el examen puede reemplazar a la peor nota de la interrogación, siempre que la nota del examen sea más alta. En otras palabras, $NIE = (I1 + I2 + I3 + Ex + Ex - \min(I1, I2, I3, Ex))/4$. No hay posibilidad de eximirse del examen.

La **nota del proyecto (NP)** corresponde al promedio ponderado de todas las entregas del proyecto. La ponderación es: Entrega 1: 5%, Entrega 2: 15%, Entrega 3: 30%, Entregas 4: 25%, Entrega 5: 25%,.

Importante: La nota de las entregas 3, 4 y 5 estarán sujetas a una evaluación de pares que puede alterar la nota del alumno. Si el profesor lo considera necesario, citará a los alumnos cuyas evaluaciones de pares indican un comportamiento irregular en las entregas.

Para aprobar el ramo, el alumno debe cumplir que NIE y NP sean mayores a 4. En ese caso, la **nota final** se calcula como $NF = (0,5NIE + 0,5NP)$. En caso contrario, $NF = \min\{NIE, NP\}$.

Fecha de las interrogaciones y fecha tentativa de entregas parciales de proyectos (este último ítem puede sufrir cambios).

Evaluaciones Escritas		Proyecto		
Evaluación	Fecha	Etapas	Enunciado	Entrega
I1	12 / 09 / 2018	I	16 / 08 / 2018	24 / 08 / 2018
I2	17 / 10 / 2018	II	29 / 08 / 2018	25 / 09 / 2018
I3	14 / 11 / 2018	III	28 / 09 / 2018	19 / 10 / 2018
Examen	05 / 12 / 2018	IV	24 / 10 / 2018	09 / 11 / 2018
		V	12 / 10 / 2018	28 / 11 / 2018

La forma de entrega de las evaluaciones del proyecto será informada a medida que avanza el curso. El plazo final es a las 23:59 del día de entrega.

Bibliografía

- Ramakrishnan, Gehrke. Database Management Systems (3ed).
- Garcia-Molina, Ullman, Widom. Database Systems: The complete book (2ed).
- Date, C. J.. Introducción a los sistemas de bases de datos (7ed).
- Elmasri, Ramez. Fundamentos de sistemas de bases de datos (5ed).
- Abiteboul, Hull, Vianu. Foundations of databases. (sobretudo para álgebra relacional y optimización).
- Phillip Greenspun. SQL for Web Nerds. <http://philip.greenspun.com/sql/>

Otros

El Departamento de Ciencias de la Computación adopta una política de tolerancia-cero frente a copias o plagios. Se sugiere revisar las políticas y penalidades que el departamento establece ante estas acciones.

El curso tiene dos canales de comunicación oficiales: Las clases y la página Web (en sistema SIDING). Se asume que toda la información que es entregada por ambos canales llega a todos los alumnos. Por lo mismo, se sugiere a los alumnos revisar la página Web constantemente.

La asistencia a clases es voluntaria, pero es responsabilidad del alumno ponerse al día con los contenidos en caso de inasistencia.