

# Sistemas de Información II

## Examen – Teoría y Práctica

**Responder brevemente, con letra legible, en hojas separadas.**

### Pregunta 1

*Mínimo 0 puntos, máximo 4 puntos: +0.5 por respuesta correcta, -0.5 por respuesta incorrecta.*

1. ¿Qué operación realiza el operador de proyección ( $\pi$ ) en álgebra relacional?
2. ¿Qué es un LEFT JOIN?
3. Indique las dos formas más típicas de codificar strings de largo variable.
4. Defina densidad de un índice.
5. ¿Qué es una función de hashing?
6. ¿En qué consiste la ley de Zipf?
7. ¿En qué se diferencia UTF-8 de ISO-8859-1 (LATIN-1)?
8. ¿Para qué sirve lematizar al construir un índice?

### Pregunta 2

*Mínimo 0 puntos, máximo 2 puntos. +0.5 por cada operación correcta, 0.0 por cada operación incorrecta.*

Considere las siguientes relaciones y escriba las operaciones que se describen en álgebra relacional.

Relación <b>R</b> (“religiones”)			Relación <b>D</b> (“denominaciones”)			
idr	nombre	adeptos	idr	idd	nombre	adeptos
1	Cristianismo	2,100,000,000	1	1	Catolicismo	1,100,000,000
2	Islam	1,500,000,000	1	2	Protestantismo	800,000,000
3	Hinduismo	781,000,000	1	3	Ortodoxia	260,000,000
4	Budismo	324,000,000	2	1	Sunismo	930,000,000
5	Sikhismo	19,000,000	2	2	Chiismo	15,000,000

1. Identificador de la religión “Budismo”
2. Nombres de las denominaciones de la religión “Cristianismo”
3. Todos los pares de nombres de religiones posibles.
4. Nombre y número de denominaciones de cada religión, incluyendo un cero para las religiones de las cuales no hay datos de denominaciones en la relación D.

## Pregunta 3

*Mínimo 0 puntos, máximo 2 puntos. +1.0 por parte correcta, 0.0 por parte incorrecta.*

1. ¿Cuántos bytes utilizaría como mínimo cada registro de la relación D? Asuma que los nombres son de largo variable. Indique los offsets de cada campo si es necesario alinear para una arquitectura de 32 bits y no se puede cambiar el orden de los campos.
2. Suponga que hay en total  $n$  registros en una tabla, y que tenemos bloques donde caben  $k$  registros por bloque, o  $m$  elementos de índice por bloque. ¿Cuántos bloques se necesitan para guardar un índice disperso? Explique por qué.

## Pregunta 4

*Mínimo 0 puntos, máximo 2 puntos. +1.0 por parte correcta, 0.0 por parte incorrecta.*

1. Construya un B-tree de orden 2 y altura 3, en que las claves son letras entre la **a** y la **z** en el orden alfabético. Asuma que el orden de inserción fue el siguiente: **a f j h k m z p**. Recuerde que en todo momento el B-tree debe tener la misma altura hacia la izquierda y hacia la derecha de la raíz. Dibuje el B-tree justo después de insertar la letra **k** y después de insertar la letra **p**.

2. Construya un índice invertido lematizado que contenga los siguientes documentos:

D1 “droga para tratar la esquizofrenia”

D2 “nueva droga para la esquizofrenia”

D3 “nuevo tratamiento para la esquizofrenia”

D4 “nuevas esperanzas para pacientes con esquizofrenia”