Organización de Datos – Curso Servetto – Evaluación de Recuperación de Textos – 11/2/09

- 1. Dada una base de 37852 documentos y un término con frecuencia global 2 que aparece e los documentos 7241 y 28234, represente la lista invertida con Golomb y justifique cómo lo hace.
- 2. Suponiendo que en el sistema de una empresa de cable se quiere recuperar el material de emisión por título en forma aproximada, y que la definición del archivo es Material ((codMaterial)i, tipo, género, título, descripción, duración, cantCapítulos), proponga la estructura de un índice que permita búsquedas de títulos que contengan un lista de términos (p.e. Lost temporada 1) o frases (p.e. "Sé lo que hicieron el verano pasado"). Describa la resolución de cada tipo de consulta.
- 3. Describa la estructura de los índices invertidos para obtener resultados de consultas ordenados por relevancia explicando el propósito de cada componente y explique cómo se resuelven las consultas.