

Львівський національний університет імені Івана Франка

Кафедра математичного моделювання соціально-економічних процесів

**ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ  
« ОСНОВИ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ »**

напрямок підготовки 6.040301 – *прикладна математика*

форма навчання: *денна*

**Виписка з навчального плану**

Форма навчання	Курс	Семестр	Загальний обсяг (год.)	Всього аудит. (год.)	у тому числі (год.):			Самостійна робота (год.)	Контрольні (модульні) роботи (шт.)	Програмні проекти (шт.)	Курсові проекти (роботи), (шт.)	Залік (сем.)	Екзамен (сем.)
					Лекції	Лабораторні	Практичні						
Денна	3	5	72	36	36			36	2			5	
		6	72	34	34			38	3				6

## 1. АНОТАЦІЯ

В основі інформаційних технологій, передусім найскладніших, якими сьогодні є системи керування базами даних, виникають такі проблеми:

- як зберігати великі об'єми інформації, забезпечуючи надійність зберігання;
- як змістовно формувати інформацію в записи і блоки;
- як ефективно досягти до інформації, враховуючи її великі об'єми.

Курс присвячено вивченню проблеми зберігання великих обсягів інформації; методів підвищення швидкодії алгоритмів обробки даних, які необхідно переміщати між оперативною пам'яттю і вторинними або третинними пристроями зберігання; методів забезпечення надійності дискових систем, які дають змогу подолати періодичні помилки читання-запису, а також зберегти дані після поломки одного чи декількох дисків; підходів до зберігання окремих елементів даних – відношень, кортежів, значень атрибутів- згідно з вимогами блочної моделі дискових даних; проблеми використання різноманітних структур даних для побудови індексів з метою забезпечення ефективного доступу до інформації; поняття багатовимірних даних, які використовують в геоінформаційних системах і системах підтримки прийняття рішень, а також особливості побудови індексів для швидкого доступу до просторових даних, зокрема сіткових файлів, роздільних геш-функцій, індексів з декількома ключами, kd-дерев, дерев квадрантів, R-дерев.

## 2. ЗМІСТ ПРОГРАМИ

### 1. Принципи зберігання інформації.

**Інформаційні технології – спільні проблеми та методи їх вирішення.** Багаторівневе предсталення даних. Бази даних як сховище інформації.

**Пристрої зберігання інформації.** Ієрархія пристроїв пам'яті. Кеш-пам'ять. Оперативна пам'ять. Віртуальна пам'ять. Вторинні пристрої зберігання. Третинні пристрої зберігання. Енергозалежні та енергонезалежні пристрої пам'яті. Диски. Внутрішні механізми дискових накопичувачів. Контролер дисків. Загальні параметри дисків. Параметри доступу до даних на диску. Записування/читання блоків. Модифікація вмісту блоків.

**Використання вторинних пристроїв зберігання.** Модель обчислень з функціями введення/виведення. Огляд алгоритмів обробки даних, які цілковито розташовані в ОП. Підвищення продуктивності систем, яким доводиться звертатися до вторинних і навіть третинних пристроїв зберігання. Особливості алгоритмів при використанні вторинних пристроїв зберігання. Сортування даних у вторинних сховищах. Сортування двофазним багатокomпонентним злиттям. Багатокomпонентне злиття і надзвичайно великі відношення.

**Методи підвищення ефективності дискових операцій.** Методи підвищення ефективності дискових систем, у яких блоки читаються і записуються у наперед відомій послідовності. Групування даних по циліндрах диска. Використання декількох дискових пристроїв. Створення дзеркальних копій дисків. Методи підвищення ефективності дискових систем, у яких блоки читаються і записуються у наперед невідомій послідовності. Впорядкування дискових операцій і алгоритм ліфта. Попереднє зчитування і крупномасштабна буферизація даних. Прийоми оптимізації дискових операцій: переваги і недоліки.

**Методи забезпечення надійності зберігання даних.** Відмови дискових пристроїв. Періодичні відмови. Контрольні суми. Стійкі сховища. Подолання наслідків помилок. Методи відновлення даних за цілковитої відмови диска. Модель відмови дискових пристроїв. Дзеркальні диски як засіб резервування. Огляд RAID-технологій. Блоки парності. Масиви RAID рівня 4,5. Відновлення даних після відмови декількох дисків.

### 2. Представлення даних на вторинних пристроях зберігання.

**Моделі даних для представлення інформації.** Огляд сучасних моделей даних. ER -модель. Об'єктно-орієнтовані моделі даних. Ієрархічна модель даних. Напівструктуровані моделі даних. Реляційна модель представлення даних. Відношення. Атрибути. Схеми. Кортежі. Домени. Ключі відношень. Представлення елементів даних на вторинних пристроях зберігання. Представлення елементів реляційних баз даних. Представлення об'єктів. Представлення елементів даних полями.

**Способи формування записів.** Записи постійної довжини. Конструювання записів постійної довжини. Заголовки записів. Групування записів постійної довжини в блоки. Елементи даних і записи змінної довжини. Записи з полями змінної довжини. Записи з полями, які повторюються. Записи змінного формату. Записи крупного об'єму. Об'єкти BLOB. Представлення адрес записів і блоків. Системи «клієнт/сервер». Логічні і структуровані адреси. Підміна покажчиків. Збереження блоків на диску з підміненими покажчиками. «Закріплені» в ОП записи і блоки. Модифікація записів. Вставляння нових записів. Вилучення записів. Оновлення записів і блоків.

### **3. Структури індексів як методи доступу до даних.**

**Індекси для послідовних файлів.** Первинні індекси для послідовних файлів. Послідовні файли. Щільні індекси. Розріджені індекси. Багаторівневі індекси. Дублікати ключових значень. Керування індексами під час модифікації даних. Вторинні індекси для послідовних файлів. Проектування вторинних індексів. Використання вторинних індексів. Додаткові рівні у вторинних індексах. Пошук документів і звернені індекси.

**Деревоподібні структури індексів.** Двійкові дерева пошуку. Довільні двійкові дерева, симетричні дерева, AVL – дерева. В – дерева та В+- дерева. Структура В-дерева та В+- дерева. Використання В-дерев. Пошук у В-деревах. Вставляння елементів у В-дерево. Вилучення елементів з В-дерева. Оцінки ефективності В-деревоподібних індексів.

**Геш-подібні структури індексів.** Статичні геш-таблиці для даних у вторинних сховищах. Особливості геш-таблиць для даних у вторинних сховищах. Вставляння записів у геш-таблицю. Вилучення записів з геш-таблиці. Оцінка ефективності гешованих індексів. Динамічні геш-таблиці для даних у вторинних сховищах. Геш-таблиці, що розширюються. Вставляння записів у геш-таблицю, що розширюється. Лінійні геш-таблиці. Вставляння записів у лінійну геш-таблицю.

### **4. Багатовимірні дані і методи їх обробки.**

**Поняття багатовимірних даних.** Інтеграція інформації з декількох джерел. Проблеми інтеграції інформації. Федеративні бази даних. Медіатори. Сховища даних. Застосування моделі багатовимірних даних. Географічні інформаційні системи. Куби даних. SQL-запити до багатовимірних даних. Запити в діапазонах значень з використанням традиційних індексів. Пошук найближчих сусідніх об'єктів з використанням традиційних індексів.

**Геш-подібні структури для багатовимірних даних.** Сіткові файли. Структура сіткового файлу. Пошук записів у сітковому файлі. Вставляння записів у сітковий файл. Оцінки ефективності структур сіткових файлів. Роздільні геш-функції. Побудова індекса за допомогою роздільних геш-функцій. Порівняння структур сіткових файлів і роздільних геш-таблиць.

**Деревоподібні структури для багатовимірних даних.** Індеси з декількома ключами. Побудова індекса з декількома ключами. Оцінки ефективності структур індексів з декількома ключами. kd-дерева. Використання kd-дерев для побудови індексів. Операції з kd-деревами. Покращення структури kd-дерева для даних у вторинних сховищах. Дерева квадрантів. Використання дерева квадрантів для побудови індексів. Операції з деревами квадрантів. R-дерева. Використання R-дерева для побудови індексів. Операції з R-деревами.

## ОСНОВНА ЛІТЕРАТУРА

1. **Ахо А., Хопкрофт Дж., Ульман Дж.** Структуры данных и алгоритмы. - М.: Издательский дом «Вильямс», 2003. – 384 с.
2. **Гарсиа-Молина Г., Ульман Дж., Уидом Дж.** Системы баз данных. Полный курс. – М.: Издательский дом «Вильямс», 2003.– 1088 с.
3. **Кнут Д.Э.** Искусство программирования. Т.3. Сортировка и поиск. – М.: Издательский дом «Вильямс», 2000. – 822 с.
4. **Копитко М.Ф.** Основи інформаційних технологій Львів: Вид. ЛНУ ім. Івана Франка, 2007.
5. **Костів О.** Структури даних. Частина 1: Тексти лекцій. – Львів: Видавничий центр ЛНУ імені Івана Франка, 2000. – 56 с.
6. **Цегелик Г.Г.** Моделювання та оптимізація доступу до інформації файлів баз даних для однопроцесорних і багатопроцесорних систем. - Львів: Вид. ЛНУ ім. Івана Франка, 2010.