

Моделі баз даних

Опрацювання даних в середовищі комп'ютерних систем передбачає визначення низки формальних, абстрактних, математичних засобів, за допомогою яких можна забезпечити відображення значень реального світу в середовище комп'ютерних інформаційних систем.

Архітектура баз даних визначає найзагальніші способи і закономірності утворення структурних одиниць, їх функції, порядок створення і застосування незалежно від їх реалізації в базах даних.

База даних – сукупність взаємопов'язаних даних, об'єднаних спільним застосуванням, середовищем зберігання та технологією опрацювання.

Моделі даних

Проблеми, пов'язані з базами даних за змістом, способами та засобами розв'язання, поділяються на такі рівні:

- задачі збереження даних;
- задачі керування даними та опрацювання даних;
- задачі відображення та сприйняття даних.

Базові принципи побудови структур даних у базах визначаються у моделі даних.

Модель даних визначають як деякий інтелектуальний засіб, що дозволяє реалізувати інтерпретацію даних відповідно до визначених вимог.

Архітектура баз даних CODASYL

Опублікована в 1971 році. Розробник Conference on Data System Languages. Передбачає такі компоненти:

- схема – глобальний опис структури бази та всіх її компонентів;
- підсхема – визначення частини бази даних з погляду кінцевого користувача чи проблеми;
- мова керування даними – інструмент опису даних, їх властивостей та дій над ними.

Недоліки такого підходу:

- висока ступінь залежності між даними та прикладними програмами;
- складність процедур навігації;
- залежність складності процедур опрацювання баз даних від їх масштабів та обсягів;
- дворівнева архітектура виключає фізичний рівень опрацювання даних.

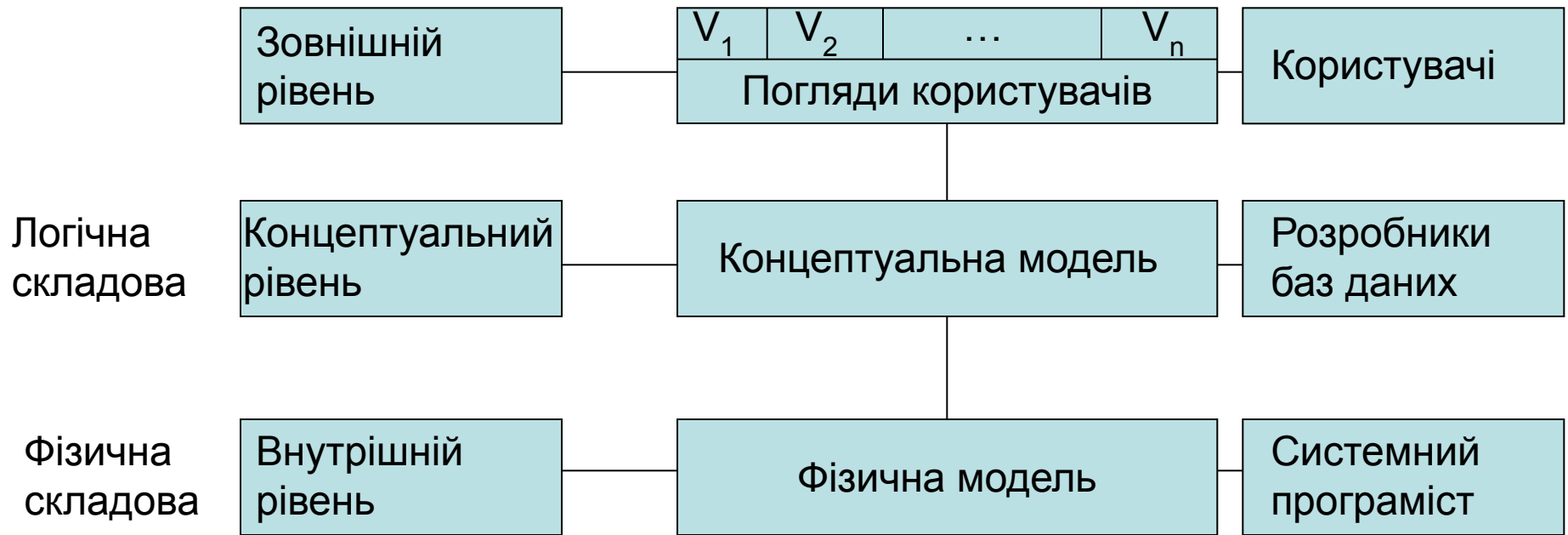
Архітектура бази даних ANSI/X3/SPARC

ANSI є подальшим розвитком концепції CODASYL. Запропонована 1975 році національним інститутом стандартів США (ANSI).

Стандарт визначає три рівні в архітектурі баз даних:

- зовнішній рівень – опис підмножини даних із застосування понять і термінів, якими оперують користувачі прикладної інформаційної системи;
- концептуальний рівень – рівень задач розробника та адміністратора баз даних, на якому описуються елементи структури баз даних, зв'язки між ними;
- внутрішній рівень - рівень системного програмування та адміністрування, на якому розв'язуються проблеми зберігання даних та доступу до них.

Архітектура бази даних ANSI/X3/SPARC



Одиниці фізичної структури даних

Фізичні моделі визначають способи утворення та впорядкування одиниць даних та методів роботи з ними на рівні зберігання.

Одиниці фізичної структури визначаються апаратними та технологічними особливостями платформи, на яких реалізована база даних.

Біт (bit) – найменша фізична одиниця даних, яка відображає один двійковий розряд.

Байт (byte) – найменша адресована одиниця; послідовність суміжних бітів, які використовуються для зберігання одного символу кодової таблиці внутрішнього коду системи (логічно завершена одиниця даних).

Слово (word) – послідовність суміжних байтів, що використовуються для зображення цілісного значення. Довжина слова визначається апаратними та системними особливостями платформи (від 2 до 8 байтів).

Фізичний запис (physical record) - послідовність суміжних байтів, що читаються з носія або записуються на носій за одну операцію вводу / виводу.

Блок, кластер (block, cluster) – це група фізичних записів, які зчитуються за одне звертання до носія. Мета блокування – зменшення кількості звернень до зовнішнього носія даних.

Фізичний файл (physical file) – множина однотипних фізичних записів, розміщених на носіях, об'єднаних спільним ідентифікатором, призначенням і місцезнаходженням. Один фізичний файл може розміщуватись на одному фізичному носії.

Каталог, папка (directory, folder) – множина фізичних файлів на носії, об'єднаних спільним призначенням, місцем збереження, ідентифікатором.

Файлова система (file system) – множина файлів, розміщених на носії даних, сумісно з засобами їх ідентифікації та адресування.

Том (volume) – один фізичний носій даних (жорсткий диск, CD, DVD, дискета тощо).

RAID (Redundant Array of Independent Disks – надлишковий масив незалежних дисків) – масив жорстких дисків, які утворюють єдиний простір збереження даних.

Способи реалізації фізичної структури даних

- Послідовна модель даних передбачає опрацювання даних в порядку їх розміщення на носії.

Перевага – простота, можливість реалізації на різних носіях.

Недолік – неможливість доступу до вибраного елемента даних без звернення до всіх попередніх.

- Індексно-послідовна модель передбачає, що для кожного опрацьованого елемента даних вводиться ідентифікатор – ключ.

Ключ наз. вбудованим, якщо він утворюється із частин самих даних, зовнішній ключ - це додаткова величина, яка не має власного застосування і вводиться виключно для ідентифікації елементів даних.

Звернення до елементів даних в індексно-послідовній моделі реалізується за значенням ключа. Для цього створюється таблиця, яка називається індекс. Індекс – відсортована таблиця, яка визначає відповідність між значенням ключа і адресою, за якою розміщений відповідний елемент даних.

- Фізична модель прямого доступу – поєднання індексно-послідовної моделі та простоти послідовної. Для ідентифікації елементів використовують їх фізичне зміщення відносно початку файлу.
- Спискова модель. Список – структура, кожний елемент якої містить посилання на інший елемент. Розрізняють списки однонаправлені та двунаправлені.

- Їєрархічна модель – узагальнення спискової. Елементи даних класифікуються за рівнями і кожному елементу надається одне або більше посилань на елементи нижнього рівня. Схеми посилань описуються деревовидним графом.
- Мережева модель дозволяє опрацьовувати дані в довільному порядку, кожен елемент може мати посилання на довільну кількість інших елементів структури. Схеми посилань описується у вигляді графу типу мережа.

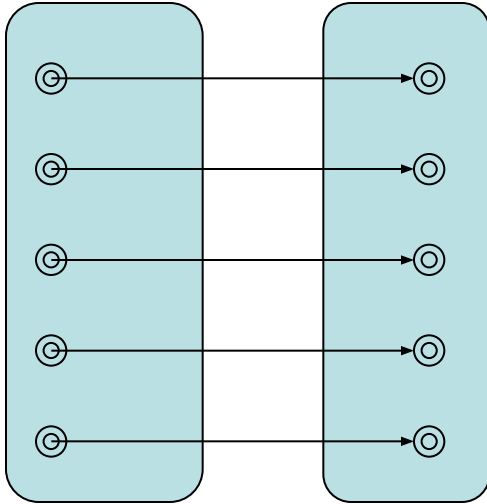
Сутність - абстракція, якою позначають будь-який об'єкт, предмет, процес, явище, що може бути однозначно ідентифікованим за множиною своїх ознак або властивостей. Один і той же самий фізичний об'єкт може виступати прототипом різних сутностей.

Атрибут – поняття, яким умовно позначають одну з властивостей сутності.

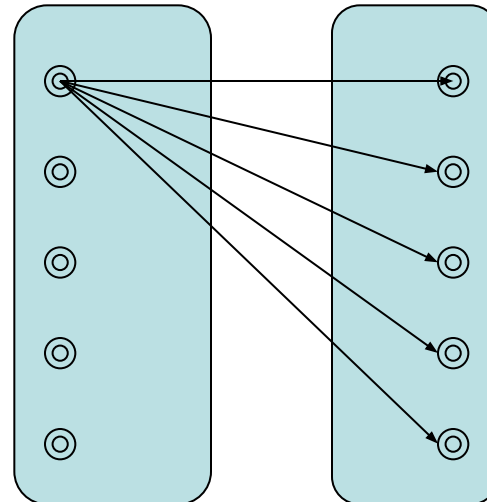
Зв'язок – стійка відповідність між примірниками сутностей. Зв'язки відображають взаємодію сутностей у певних процесах. Вони забезпечують цілісне узгоджене застосування різних частин бази даних.

Види та властивості зв'язків

Зв'язок типу 1:1 – "один – до - одного"

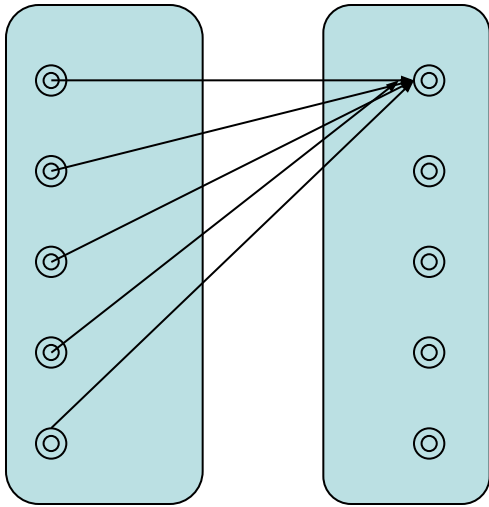


Зв'язок типу 1:N – "один – до - багатьох"

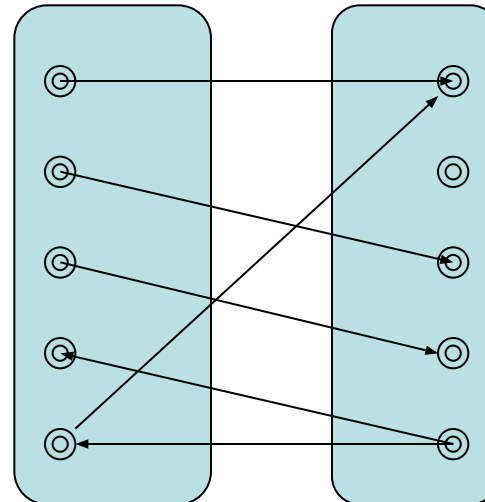


Види та властивості зв'язків

Зв'язок типу N:1 – "багато – до - одного"

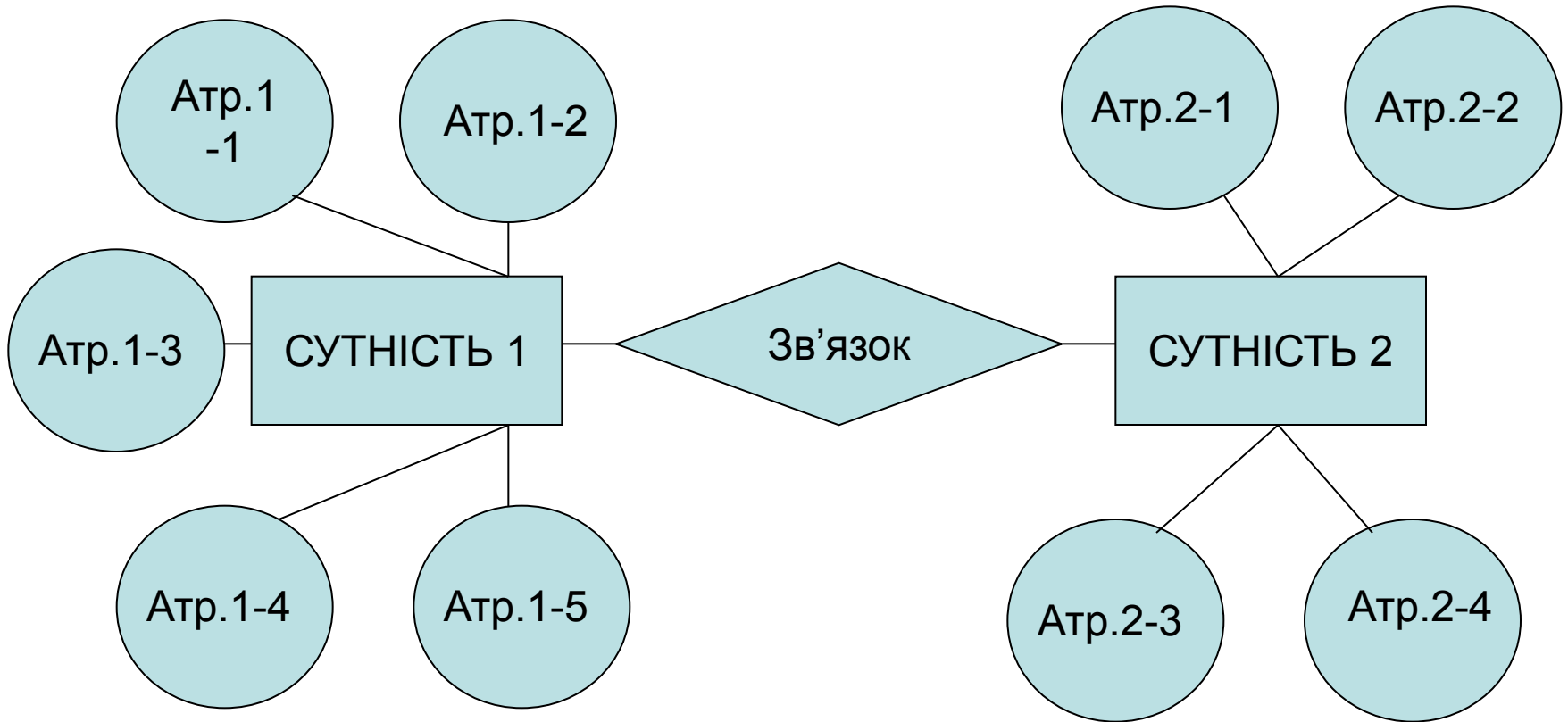


Зв'язок типу N:M – "багато – до - багатьох"



ER-діаграма

Елементи діаграми “ сутність – зв’язок ” у нотації Чена



Атрибути класу сутностей “Студент”

The diagram illustrates the attributes of the 'Student' class. A table lists specific student records, and a vertical bar to the right represents the collection of attribute values. An arrow points from the 'Група' column to the vertical bar, and another arrow points from the vertical bar to the 'Значення атрибутів' label.

№ студ. квитка	Прізвище	Ім'я	По- батькові	Група	Спеціальність	Факультет	Кафедра
111111	Петренко	Іван	Петрович	АВ-21	АУТП	КНІТ	АУВП
221122	Іваненко	Сергій	Хомович	АВ-21	АУТП	КНІТ	АУВП
121212	Седір	Микола	Хомич	АВ-21	АУТП	КНІТ	АУВП
...

Значення атрибутів