

Bazy danych - notatki z wykładu

Jakub Kwaśny, 2025

Wstęp

Dokument zawiera streszczenie w formie notatek najważniejszych pojęć pojawiających się na wykładzie.

Rozdział 1. Kiedyś to były bazy...?

Dane – symbole lub zbiory symboli odzwierciedlające pewne fakty.

Baza danych – zbiór powiązanych logicznie danych wraz z definicją ich struktury.

Systemy oparte na przetwarzaniu plików – architektura sprzed ery systemów zarządzania bazami danych. Polega na tym, że każda aplikacja ma swoje prywatne pliki, w których przechowuje potrzebne jej dane. Wady takiego podejścia to m.in. powielanie danych oraz potencjalne niespójności, zależność kodu programu od sposobu przechowywania danych oraz ograniczenia w raportowaniu.

System zarządzania bazą danych (SZBD) – oprogramowanie działające jako pośrednik w dostępie do bazy danych. Obsługuje odczyt i modyfikację danych przez aplikację, zwykle za pomocą języka SQL. Kontroluje powielanie danych oraz ich spójność, odpowiada za bezpieczeństwo i standaryzację. Umożliwia uzyskanie większej ilości informacji dzięki łączeniu danych z wielu źródeł, a także możliwości bezpośredniego odpytania przez użytkowników z pominięciem aplikacji. Wadami SZBD jest większa złożoność systemów, dodatkowe koszty oprogramowania, sprzętu oraz administrowania. Przykładowe SZBD: Oracle Database, MySQL, SQL Server, MongoDB.

Architektura trójwarstwowa – koncepcja zakładająca możliwość analizy bazy danych na poziomie zewnętrznym (to, co widzi użytkownik mający dostęp do bazy), konceptualnym (układ danych w tabelach i zależności między nimi) i wewnętrznym (fizyczna realizacja przechowywania danych). Zmiany na niższych poziomach, czyli konceptualnym (np. dodanie nowej kolumny, rozdzielenie tabeli na dwie) lub wewnętrznym (np. przeniesienie bazy na szybszy dysk) nie powinny powodować konieczności zmian na wyższych poziomach.

Architektura klient-serwer – koncepcja zakładająca podział aplikacji na część działającą na serwerze aplikacyjny (przetwarzającą dane) oraz część działającą na urządzeniu użytkownika (klient obsługujący interfejs, wstępnie sprawdzający poprawność danych, wyświetlający wyniki). Odciąża to centralny serwer od niektórych zadań, umożliwia większą skalowalność. Jako etap pośredni często wykorzystywany jest tzw. load balancer, czyli oprogramowanie rozdzielające zapytania do serwera na kilka serwerów aplikacyjnych, umożliwiające równomierne rozłożenie ruchu pomiędzy te serwery i zapewniające dostępność w przypadku awarii danego serwera aplikacyjnego. Często skonfigurowana jest również tzw. pula połączeń, polegająca na utrzymaniu stałej liczby połączeń do bazy danych zamiast łączenia się przy okazji każdego zapytania ze strony klienta.

Rozdział 2. Jak matematycy patrzą na bazy danych?

Encja – reprezentacja wyobrażonego lub rzeczywistego obiektu. Nazwami encji są rzeczowniki w liczbie pojedynczej. Pomiedzy encjami mogą występować relacje (np. Pracownik obsługuje Zamówienie). Encje mają cechy nazywane **atributami**. Encja w bazie danych powiązana jest z tabelą, a atrybuty to kolumny tej tabeli. Wiersze tabeli nazywamy czasami **krotkami**.

Model relacyjny – model bazy danych, w którym tabele traktowane są jako relacje matematyczne, czyli podzbiory iloczynu kartezjańskiego **diedziny** (zbiorów wszystkich możliwych wartości danego

atrybutu). Wynikają z tego pewne własności, np.: kolejność wierszy w tabeli jest nieistotna, ale kolejność kolumn jest istotna; wiersze nie powtarzają się. **Stopień relacji** to liczba kolumn tabeli, a **moc relacji** to liczba wierszy.

Nadklucz – atrybut lub zbiór atrybutów które pozwalają jednoznacznie zidentyfikować krotkę. Minimalny nadklucz nazywamy **kluczem kandydującym**. W modelu relacyjnym dla każdej tabeli musimy wybrać i wyróżnić pewien klucz kandydujący. Taki wybrany klucz kandydujący nazywamy **kluczem głównym** tabeli.

Klucz obcy – atrybut lub zbiór atrybutów, który odpowiada kluczowi głównemu pewnej (być może tej samej) relacji.

Język SQL (Structured Query Language) – ustandaryzowany język do komunikacji z relacyjnymi bazami danych. Składa się z: Data Definition Language (tworzenie, modyfikacja i usuwanie obiektów w bazie danych, np. tabel), Data Manipulation Language (wstawianie, modyfikacja i usuwanie wierszy z tabel), Data Query Language (pozyskiwanie danych z tabel; czasami uznawany za podzbiór DML) oraz Data Control Language (zarządzanie dostępem do danych, np. przydzielanie uprawnień czy zmiana hasła).