**Wykład 1. Wprowadzenie do problematyki baz danych**

1. Funkcjonalność systemu zarządzania bazą danych (SZBD) obejmuje:

- Struktury danych

- Mechanizmy współbieżnego dostępu do danych

- Mechanizmy optymalizacji dostępu do danych

- Język bazy danych

1. Model danych definiuje:

- Struktury danych, operacje na danych, ograniczenia integralnościowe

1. Ze względu na model danych wyróżnia się bazy danych:

- Semistrukturalne

- Relacyjne

1. Zadaniem analityków systemowych kest:

- Analiza wymagań projektowanego systemu informatycznego

1. Architektura odniesienia systemu bazy danych wg ANSI/SPARC zakłada:

- Istnienie schematu wewnętrznego, implementacyjnego i zewnętrznego

1. Spójność bazy danych oznacza:

- Jej odporność na anomalie będące wynikiem współbieżności dostępu do baz danych

- Jej poprawność z punktu widzenia zdefiniowanych reguł

1. Wybrać wszystkie cechy charakteryzujące współbieżny dostęp do danych:

- Konfliktowe operacje równoczesnego odczytu i zapisu tych samych danych

- Dostęp do wielu danych jednocześnie przez wielu różnych użytkowników

- Konfliktowe operacje równoczesnego zapisu tych samych danych

1. Wskazać system, który nie jest komercyjnym SZBD

- PostgreSQL

1. Metadane to:

- Dane opisujące bazę danych

1. Języki 4GL projektowania aplikacji baz danych:

- Umożliwiają bezpośrednie zagnieżdżenie poleceń SQL w kodzie

**Wykład 2. Relacyjny model danych**

1. Relacyjny model danych definiuje następujące operacje:

- Selekcję, projekcję, połączenie, operacje na zbiorach

1. Relacją nazywamy:

- Podzbiór iloczynu kartezjańskiego dziedzin jej atrybutów

1. Baza danych jest:

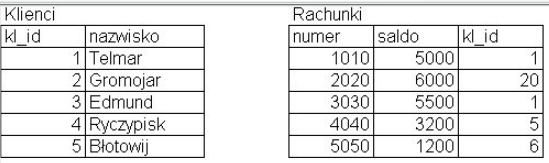
- Zbiorem relacji

1. Kluczem podstawowym relacji jest:

- Wybrany zbiór atrybutów, których wartości jednoznacznie identyfikują każdą krotkę relacji

1. Dane są dwie relacje: A = {a, b, c, d, e} i B = {c, d, f, g, h}. Wynikiem operacji A SUMA B, A MINUS B, A ILOCZYN B są:

- A SUMA B = {a, b, c, d, e, f, g, h}; A MINUS B = {a, b, e}; A ILOCZYN B = {c, d}

1. Dane są dwie relacje: Klienci i Rachunki. Wynikiem operacji Klienci\*Rachunki jest: 

- Relacja {<1, ‘Telmar’, 1010, 5000>, <1, ‘Telmar’, 3030, 5500>, <5, ‘Błotowij’, 5050, 1200>}

1. NULL oznacza:

- Wartość pustą/nieokreśloną

1. Wskazać własność, która nie jest własnością relacji:

- Porządek atrybutów w relacji jest określany w momencie tworzenia relacji

1. Ograniczenie intergralnościowe zawężające dziedzinę definiuje:

- Zakres dozwolonych wartości dla atrybutu

- Zbiór dozwolonych własności dla atrybutu

1. Dana jest relacja PakietyInternetowe{pakietID, przepustowość, cena} z dwoma następującymi krotkami {

- <2, ‘Smętek’, 3>

**Wykład 3. Modelowanie danych: Model związków-encji**

1. Encja jest:

- Obiektem konceptualnego modelu związków-encji

1. Związek 1:1 obustronnie opcjonalny pomiędzy encjami A i B oznacza, że:

- Jedno wystąpienie encji B może wchodzić w związek z dokładnie jednym wystąpieniem encji A

- Jedno wystąpienie encji A może wchodzić w związek dokładnie z jednym wystąpieniem encji B

1. Model implementacyjny służy do:

- Określenia struktur danych służących do przechowywania danych

1. Każda encja posiada następujące cechy:

- Nazwę i atrybuty

1. Typ asocjacji oznacza:

- Liczbę wystąpień encji, która może wchodzić w związek z określoną liczbą wystąpień innej encji

1. Związek 1:M obustronnie obowiązkowy pomiędzy encjami A i B oznacza, że:

- Jedno wystąpienie encji A musi wchodzić w związek z przynajmniej jednym wystąpieniem encji B

- Jedno wystąpienie encji B musi wchodzić w związek dokładnie z jednym wystąpieniem encji A

1. Stopnień związku określa:

- Z iloma encjami wchodzi w związek dana encja

1. Encja słaba to:

- Nie ma własnego identyfikatora

1. Związki wyłączone:

- Oznacza, że wystąpienie encji może wchodzić tylko w jeden związek ze zbioru związków wyłączonych

- Oznacza się na diagramie w notacji Barkera za pomocą łuku

1. Semantyka hierarchii jest następująca:

- Wystąpienie podencji jest jednocześnie wystąpieniem nadencji

**Wykład 4. Transformacja modelu ER do modelu relacyjnego**

1. Nazwa tabeli powstającej z encji jest:

- Nazwą encji (rzeczownikiem) w liczbie mnogiej

1. Klucz obcy:

- Jest zbiorem atrybutów wskazujących na klucz podstawowy tej samej lub innej tabeli

1. Identyfikator encji transformuje się do:

- Klucza podstawowego

1. Związek 1:1 pomiędzy encjami A i B, obowiązkowy od strony encji B transformuje się do:

- Klucza obcego w tabeli B; klucz ten ma ograniczenie NOT NULL

1. Związek 1:M obustronnie opcjonalny pomiędzy encjami A (jeden) i B(wiele) transformuje się do:

- Klucza obcego w tabeli B

1. Związek 1:M obowiązkowy od strony wiele pomiędzy encjami A (jeden) i B (wiele) transformuje się do:

- Klucza obcego w tabeli B z ograniczeniem integralnościowym NOT NULL dla tego klucza obcego

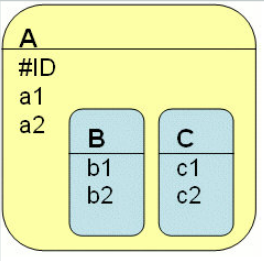
1. Związek M:N obustronnie opcjonalny pomiędzy encjami A i B transformuje się do:

- Tabeli pośredniej z dwoma kluczami obcymi, oba z ograniczeniem integralnościowym NOT NULL

1. Schemat pierwszy transformacji hierarchii generalizacji daje w wyniku jedną tabelę z dodatkowym atrybutem nie wynikającym z atrybutów nadencji i podencji. Dziedzina wartości tego dodatkowego atrybutu:

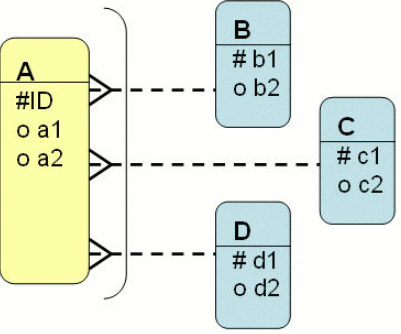
- Zawiera tyle wartości ile jest podencji

1. Dana jest hierarchia encji jak na rysunku. Transformacja tej hierarchii wg schematu 2 da w rezultacie (pominięto ograniczenia integralnościowe):



- 2 tabele: B{ID, a1, a2, b1, b2}, C{ID, a1, a2, c1, c2}

1. Dany jest model ER jak na rysunku. Transformacja encji A i związków ma postać:



- A{ID PRIMARY KEY, a1, a2, b1 NULL FOREIGN KEY, c1 NULL FOREIGN KEY, d1 NULL FOREIGN KEY}

**Wykład 5. Normalizacja schematów logicznych relacji**

1. Dana jest relacja R = {A, B, C, D} zawierająca następujące kroki

- A -> D i C -> D i A -> C

1. Normalizacja służy:

- Zdekomponowaniu relacji na fragmenty pozbawione anomalii wstawiania, modyfikowania i usuwania

1. Atrybutem podstawowym jest:

- Atrybut, który stanowi część klucza relacji

1. Druga postać normalna:

- Gwarantuje, że wartości wszystkich atrybutów są atomowe

- Gwarantuje, że wszystkie atrybuty wtórne są w pełni funkcyjnie zależne od klucza relacji

1. Trzecia postać normalna gwarantuje:

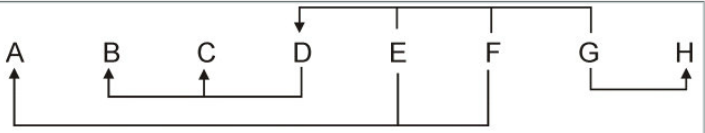
- Że w relacji nie ma przechodnich zależności funkcyjnych atrybutów wtórnych od klucza relacji

- Że wszystkie atrybuty wtórne są w pełni funkcyjnie zależne od klucza relacji

- Że relacja jest w drugiej postaci normalnej

- Że wartości wszystkich atrybutów są atomowe

1. Dana jest tabela T (A, B, C, D, E, F, G, H) i zależności funkcyjne przedstawione na rysunku. Określić w której postaci normalnej znajduje się tabela T.



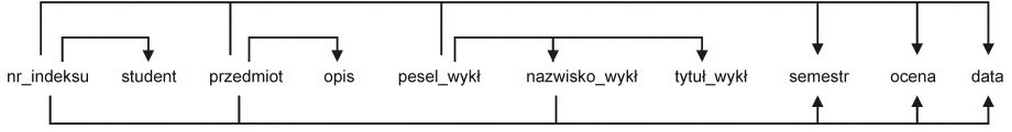
- 2NF

1. Zależność wielowartościowa A ->-> B pomiędzy atrybutem A i B relacji R = {A, B, C}:

- Określa związek pomiędzy atrybutami A i B

- Określa niezależność atrybutów B i C od siebie

1. Dane są zależności funkcyjne pomiędzy atrybutami relacji, jak na rysunku. Transformacja tej relacji do 3NF da w wyniku następujące relacje:



- Oceny = {nr\_indeksu, przedmiot, pesel\_wykł, semest, ocena, data}, Studenci = {nr\_indeksu, student}, Przedmioty = {przedmiot, opis}, Wykladowcy = {pesel\_wykl, nazwisko\_wykl, tytuł\_wykł}

1. Dana jest relacja R = {X, Y, Z}, gdzie X->->Y. Zawartość relacji jest następująca:

- <a, c, e>

1. Postać BCNF

- Stanowi warunek dostateczny, ale nie konieczny na to, aby relacja była w 3NF.

**Wykład 6. Organizacja plików**

1. Organizacja plików określa:

- Sposób uporządkowania rekordów w plikach dyskowych

1. Współczynnik blokowania określa:

- Maksymalną liczbę rekordów, które w całości mieszczą się w bloku danych

1. Wolne miejsce pozostawione w każdym bloku danych jest przeznaczone na:

- Rozszerzanie się rekordów na skutek ich modyfikowania

1. Wskazać wszystkie własności, które nie obowiązują dla pliku nieuporządkowanego:

- Rekordy mają zawsze zmienny rozmiar

- Rekordy są wstawiane w wolne szczeliny w całym pliku

- Rekordy mają zawsze stały rozmiar

- Rekordy są uporządkowane według pola niekluczowego

1. Maksymalny koszt dostępu do pliku nieuporządkowanego przy wyszukiwaniu rekordów jest określony wzorem:

- N(D+RC)

1. Algorytmem przeszukiwania pliku uporządkowanego z kryterium nałożonym na pole porządkujące jest:

- Algorytm połowienia binarnego

1. Wynikiem działania funkcji haszowej 111 MOD 10 jest

- 1

1. Wskazać wszystkie poprawne cechy opisujące kolizję:

- Jest spowodowana funkcją haszową, która dla różnych wartości atrybutu haszowego generuje te same numery szczelin

- Polega na tym, że dwa lub więcej rekordów trafiają do tej samej szczeliny

1. Jedną z metod rozwiązywania kolizji jest:

- Łańcuchowanie

1. Podstawowa różnica pomiędzy haszowaniem wewnętrznym, a zewnętrznym jest następująca:

- Haszowanie wewn. adresuje szczeliny, a zewn. – logiczne obszary dyskowe

**Wykład 7. Indeksy**

1. Wybrać odpowiedź, która charakteryzuje indeks rzadki:

- Liczba rekordów indeksu jest mniejsza niż liczba rekordów danych

1. Wskazać wszystkie cechy opisujące indeks wtórny:

- Założony na atrybucie nieporządkującym nieuniknionym

- Jest indeksem gęstym

- Założony na atrybucie nieporządkującym unikalnym

1. W systemie pojawia się dziennie średnio 80% zapytać z warunkiem miasto = ‘Poznań’, 85% zapytań z warunkiem wykształcenie = ‘wyższe’ i miasto = ‘Poznań’, 10% zapytań z warunkiem województwo = ‘Wielkopolskie’

Jaki zbiór indeksów będzie najwłaściwszy do optymalizacji powyższego zbioru zapytań?

- Indeks <miasto, wykształcenie> i indeks <województwo>

1. Wybrać odpowiedzi, które charakteryzują indeks ISAM:

- Statyczny

-Dwupoziomowy

- Struktura zależna od sprzętu

1. Zrównoważenie indeksu B+-drzewo oznacza, że:

- Liczba poziomów liczona od korzenia do dowolnego liścia jest taka sama

1. Rząd drzewa określa:

- Maksymalną liczbę wskaźników w korzeniu

- Maksymalną liczbę wskaźników w węzłach wewnętrznych

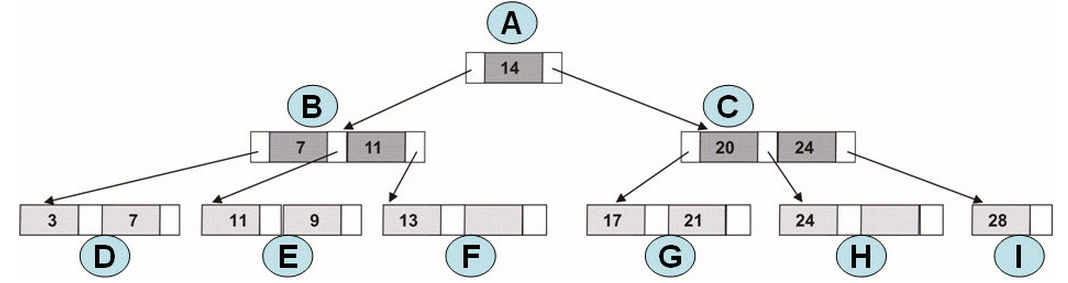
1. Niech p oznacza rząd liścia. Minimalna liczba wartości klucza w liściu jest określona jako:

- Największa liczba całkowita mniejsza lub równa p/2

1. Obliczyć wysokość indeksu B+-drzewo i liczbę odczytów bloków indeksu do wyszukania pojedynczego rekordu, przy następujących założeniach: liczba indeksowanych rekordów r=500000; rozmiar jednego rekordu R=100B; rozmiar bloku dyskowego BL=2048B; rozmiar wskaźnika P=6B; rozmiar wartości klucza indeksu K=8B.

- p=146, liczba odczytów indeksu=3

1. Zidentyfikować błędy w strukturze indeksu B+-drzewo z rysunku



- E: zły porządek kluczy; F: brak klucza o wartości 14; G: klucz 21 do liścia H; G: zamiast klucza 21 – klucz 20.

1. Dany jest indeks B+- drzewo jak na rysunku. Jaka będzie zawartość węzła B po wstawieniu do indeksu wartości 5.

- Węzeł B rozszczepi się na węzeł z wartością 5 i węzeł z wartością 11.

**Wykład 8. Przetwarzanie transakcyjne**

1. Cele stosowania transakcji w bazach danych są następujące:

- Zapewnienie właściwej efektywności i wykonywania zapytać

- Zapewnienie spójności danych w przypadku awarii sprzętowo-programowych

- Zapewnienie spójności danych w sytuacji współbieżnego dostępu do tych samych danych

1. Transakcja w bazie danych jest:

- Zbiorem poleceń, które są wykonywane jako spójna jednostka

- Jednostką interakcji użytkownika z bazą danych

1. Cecha atomowości transakcji oznacza, że:

- Transakcja w całości musi zostać zatwierdzona lub wycofana

1. Cecha trwałości transakcji oznacza, że:

- Dane wprowadzone do bazy danych przez zatwierdzoną transakcję są trwałe

1. Cecha spójności transakcji oznacza, że

- Stan bazy danych po zakończeniu transakcji jest spójny, czyli poprawny

1. Cecha izolacji transakcji oznacza, że:

- Transakcje wzajemnie nie widzą zmian w danych wprowadzonych przez niezatwierdzone transakcje

1. Commit jest:

- Poleceniem zakończenia transakcji z jej zatwierdzeniem

1. W realizacji sekwencyjnej transakcji T1 i T2, takich że T1 poprzedza T2:

- Wszystkie operacje T1 poprzedzają operacje T2

1. Realizacja zbioru równocześnie działających transakcji T1, T2…, Tn jest uszeregowalna jeżeli:

- Jest ona równoważna dowolnej sekwencyjnej realizacji zbioru transakcji

1. Graf uszeregowalności:

- Jest wykorzystywany do określania uszeregowalności zbioru transakcji

- Może być acykliczny

**Wykład 9. Algorytmy zarządzania współbieżnym wykonywaniem transakcji**

1. Z punktu widzenia blokowania, dana X może znaleźć się w jednym z następujących stanów:

- Zablokowana do odczytu, zablokowana do zapisu, niezablokowana

1. Wskazać blokady, które nie są kompatybilne:

- W z R

- R z W

- W z W

1. Transakcja T realizuje sekwencyjnie operację odczytu danej X, a następnie modyfikacji tej danej. W tym celu transakcja T:

- Zakłada blokadę R na danej X, odczytuję daną, zakłada blokadę W na danej X i zapisuje X na dysk

1. Algorytm 2PL:

- Zarządza zakładaniem i zwalnianiem blokad

- Jest stosowany zarówno w zatwierdzaniu jak i wycofywaniu transakcji

- Składa się z fazy ekspansji i fazy zwijania

1. Dane są dwie transakcje T1 i T2 takie, że TS(T1)

- T2 zostanie wycofana, T1 uzyska blokadę R na X, T2 zostanie uruchomiona z poprzednim znacznikiem czasowym

1. Dane są dwie transakcje T1 i T2 takie, że TS(T1)

- T2 będzie realizowana, T1 będzie wycofywana i wznawiana z pewnym opóźnieniem do momentu zakończenia się T2

1. Graf oczekiwania:

- Służy do wykrywania zakleszczeń

1. Dane są trzy transakcje: T1, T2 i T3 realizowane współbieżnie, jak pokazano na rysunku. Czy taka realizacja prowadzi do zakleszczenia? *(brak rysunku ☹)*

- Tak, pomiędzy T1, T2 i T3

[**Wykład 10. Algorytmy zarządzania współbieżnym wykonywaniem transakcji**](http://wazniak.mimuw.edu.pl/index.php?title=Bazy_danych/Wyk%C5%82ad_10)

1. W blokowaniu hierarchicznym:

- Zablokowanie obiektu podrzędnego wymaga uprzedniego jego obiektu nadrzędnego

- Blokowaniu podlegają m.in. tablice i rekordy

- Założenie blokady na obiekcie nadrzędnym oznacza zablokowanie wszystkich jego obiektów podrzędnych

1. Blokada RIW założona na tabeli TAB oznacza:

- Blokowanie wszystkich rekordów tabeli Tab w trybie „do odczytu”, przy czym niektóre rekordy są blokowane w trybie „do zapisu”

1. W algorytmie znaczników czasowych:

- Transakcja T może modyfikować tylko te dane, których znaczniki czasowe są mniejsze niż znacznik czasowy T

1. Transakcja w trybie READ ONLY:

- Umożliwia modyfikowanie danych, których znaczniki czasowe są mniejsze od znacznika czasowego transakcji modyfikującej

1. Poziom izolacji READ COMMITED:

- Zapewnia, że każda transakcja operuje wyłącznie na danych zatwierdzonych przez inne transakcje

1. Poziom izolacji transakcji określa:

- Jakie operacje można wykonać w ramach transakcji i jakie dane „widzi” transakcja

**Wykład 11. Recovery – Transakcyjne odtwarzanie bazy danych po awarii**

1. Miękki model awarii to:

- Awaria oprogramowania

1. Wskazać elementy architektury systemu, które nie należą do architektury odtwarzania po awarii:

- Transakcja

- Optymalizator poleceń SQL

- Aplikacje użytkownika

1. Force() jest operacją:

- Zapisania rekordów logu z bufora na dysk

1. LRU to:

- Algorytm zarządzający stronami w buforze danych

1. Rekord logu:

- Jest wykorzystywany w czasie odtwarzania bazy danych po awarii twardej

- Jest tworzony i zapisywany do bufora logu dla każdej operacji aktualizacji obiektu bazy danych i operacji inicjacji transakcji

- Jest wykorzystywany w czasie odtwarzania bazy danych po awarii miękkiej

1. Dane są trzy rekordy X=1, Y=2, Z=3. Dana jest następująca sekwencja operacji realizowanych przez transakcje T1 i T2: T1:r(X, 1), T2:r(Y, 2), T1:w(X, 10), T1:r(Z, 3), T1:w(Z, 30), T2:W(Y, 20), ‘awaria’. W momencie oznaczonym jako ‘awaria’ następuje twarda awaria. Jaki będzie stan bazy danych (wartości rekordów X, Y, Z) po jej odtworzeniu?

- X=1, Y=2, Z=3

1. W fazie ROLLBACK:

- Dane wprowadzone przez niezatwierdzone transakcje są wycofywane

1. Strategia WAL zapewnia, że:

- Strona bufora danych nie zostanie zapisana na dysku przez zapisaniem rekordu logu

1. Punkt kontrolny umożliwia:

- Wycofanie lub zatwierdzenie fragmentu transakcji

**Wykład 12. Optymalizacja zapytań**

1. Zapytanie jest przetwarzane w następujących krokach:

- Dekompozycja -> optymalizacja -> generowanie kodu -> wykonanie

1. *(treść niepełna)*
2. *(treść niepełna)*
3. Dane jest wyrażenie algebry relacji jak na rysunku. Jest ono równoważne następującemu wyrażeniu *(brak rysunku)*

- (SIGMA c (R JOIS S)) SUMA (SIGMA c (R JOIN T))

1. *(treść niepełna)*

[**Wykład 13. Optymalizacja zapytań**](http://wazniak.mimuw.edu.pl/index.php?title=Bazy_danych/Wyk%C5%82ad_13)

1. Modułem odpowiedzialnym za zbudowanie planu zapytania jest:

- Optymalizator zapytań

1. Drzewo zapytania/plan zapytania:

- Reprezentuje kolejne operacje elementarne wykonywane przez optymalizator zapytań w celu wykonania zapytania

1. W optymalizacji kosztowej każde zapytanie opisane jest:

- Kosztem wykonania każdej operacji elementarnej w planie i szacunkowym rozmiarem wyniku otrzymanego poprzez wykonanie każdej operacji elementarnej

1. Informacja o liczbie różnych wartości atrybutu jest przechowywana:

- W katalogu/słowniku bazy danych

1. Dana jest relacja R zawierająca 10000 kortek. Relacja ta posiada atrybut A, który może przyjąć 5 różnych wartości. Rozkład wartości w krotkach jest równomierny. Selektywność atrybutu A wynosi:

-0,2

1. Dane są trzy relacje R1, R2 i R3, które posiadają odpowiednio atrybuty A, B i C. Każdy z tych atrybutów ma równomierny rozkład swoich wartości. A i B mogą przyjąć jeden z czterech różnych wartości, a C może przyjąć jedną z dwóch różnych wartości. Zbiory wartości atrybutów A, B i C są rozłączne. Każda z relacji zawiera 1000 krotek. Jaka będzie kardynalność wyniku następującego zapytania, przy założeniu, że operacja UNION (suma zbiorów) nie eliminuje duplukatów:

Select \* from R1 where A=’a1’

UNION

Select \* from R2 where B=’b1’

UNION

Select \* from R3 where C=’c1’;

-1000

1. Histogram opisuje:

-Rozkład wartości atrybutu

1. Optymalizator zapytań jest charakteryzowany przez:

- Algorytm przeszukiwania przestrzeni stanów

- Reguły transformacji

- Funkcję kosztu