

Egzam SBD

Zestaw 1

//stworzone przy użyciu OCR, więc mogą się pojawić krzaki w tekście

Zad. 1. Opracuj schemat EER dla systemu bankowego. Wynik analizy mini-świata jest następujący: Bank prowadzi rachunki terminowe i ROR dla osób fizycznych i osób prawnych. Klient-osoba prawna jest opisany przez: nazwę, NIP, adres, nazwisko i numer telefonu osoby reprezentującej daną instytucję. Klient-osoba fizyczna jest opisany przez: imię, nazwisko i adres. Każdy klient, zarówno osoba fizyczna, jak i prawna posiada zbiór numerów kontaktowych. Rachunek bankowy posiada swój numer, aktualne saldo, datę ważności salda i walutę rachunku. Z każdym rachunkiem jest związana historia operacji, zawierająca: kolejny numer operacji, datę operacji, rodzaj operacji (np. wpłata, wypłata, odsetki, przelew) i kwotę operacji. Numery operacji są nadawane od 1 do n, niezależnie dla każdego rachunku. Każdy klient może posiadać wiele rachunków, ale ten sam rachunek jest własnością wyłącznie jednego klienta. Do danego rachunku można jednak udzielić pełnomocnictwa wielu osobom. Pełnomocnik jest opisany przez: imię, nazwisko i adres. Przedstaw diagram EER odpowiadający powyższemu opisowi miniświata, Określ i zaznacz w modelu identyfikatory encji.

Zad. 2. Przetransformuj poniższy diagram encji-związków (EER) do schematu relacyjnego. Zaznacz klucze główne oraz klucze obce w schematach relacji, oraz, dodatkowo, zaznacz atrybuty z wartościami obowiązkowymi i opcjonalnymi.

Zad. 3. Dany jest następujący schemat pojęciowy. Studenci realizują projekty zgodnie z przydziałem. Każdy student studiuje na określonym kierunku.

Schemat bazy danych obejmuje 4 poniższe relacje:

KIERUNKI

K_SYMBOL varchar2(3) primary key K_NAZWA varchar2(20) not null

STUDENCI

S_NR_INDEKSU number(5,0) primary key

S_NAZWISKO varchar2(20) not null

S_IMIE varchar2(20) not null

S_SREDNIA number(3,2) not null

S_STYPENDIUM number(6,2)

S_K_SYMBOL varchar2(3) not null references KIERUNKI(K_SYMBOL)

PROJEKTY

P_ID_PROJEKTU number(5,0) primary key P_NAZWA varchar2(20) not null

P_OPIS varchar2(500)

PRZYDZIAŁY

PR_P_ID_PROJEKTU number(5,0) references PROJEKTY(P_ID_PROJEKTU)

PR_S_NR_INDEKSU number(5,0) references STUDENCI(S_NR_INDEKSU) primary key (PR_P_ID_PROJEKTU, PR_S_NR_INDEKSU)

Sformułuj w języku SQL następujące zapytania wyrażone poniżej w języku naturalnym.

a) Wyświetl symbole i nazwy kierunków oraz liczbę studentów studiujących na każdym z kierunków. Uwzględnij tylko kierunki na których studiuje więcej niż 15 studentów. Wyniki uporządkuj malejąco wg liczby studentów.

b) Wyświetl numery indeksów studentów których średnia ocen przekracza średnią dla ich kierunku. Pomiędzy studentów przydzielonych do projektu o nazwie 'ALGORYTMY' uwzględnij ich przy wyznaczaniu średniej dla kierunku).

c) Przyznaj stypendium w wysokości 300 studentowi (studentom) o najwyższej średniej spośród studentów, którzy nie otrzymują stypendium.

d) Usuń projekty, do których nie przydzielono studentów, a których nazwa zaczyna się na literę 'A' lub kończy się na literę 'Y'.

Zad. 4. Dana jest relacja R przedstawiona poniżej.

R		
A	B	C
a1	b1	c1
a1	b2	c2
a1	b2	c1
a1	b1	c2

R1	
A	B
a1	b1
a1	b2

R2	
A	C
a1	c2
a1	c1

Relacja R została zdekomponowana na dwie relacje R1 i R2. takie, że $R1(A, B)$. $R2(A, C)$.

1. Czy połączenie naturalne relacji R1 i R2 daje w wyniku oryginalną relację R?
2. Czy dekompozycja relacji R na relacje R1 i R2 jest dekompozycja bez utraty informacji? Uzasadnij odpowiedź.

Zad. 5. Dany jest schemat relacji $R = (A, B, C, D, E, F, G, H)$. Dany jest również następujący zbiór zależności funkcyjnych dla schematu R:

$A \rightarrow B, C$

$ABE \twoheadrightarrow CDGH$

$C \twoheadrightarrow G, D$

$D \rightarrow G, E \rightarrow F$

1. Podaj klucz (klucze) schematu R.
2. Schemat relacji R jest w pierwszej postaci normalnej. Zdekomponuj schemat R do zbioru schematów w 2NF. ale nie 3NF. Uzasadnij, że otrzymane schematy są w 2NF i że nie są w 3NF.
3. Zdekomponuj otrzymany zbiór schematów w kroku 2 do zbioru schematów w 3NF. Czy otrzymane w wyniku dekompozycji schematy są również w BCNF?

Zad. 6. Na kluczu podstawowym, który jest kluczem sortowania pliku danych, założono indeks B+ drzewo. Klucz **podstawowy** przybiera wartości od 1 do 10.000.000. Oblicz, zakładając równomierny rozkład wartości indeksowanego atrybutu, ile średnio należy wykonać operacji I/O w celu odczytania, z wykorzystaniem indeksu, wszystkich rekordów pliku z wartością klucza z przedziału domkniętego [2.000.000, 2.999.999].

Dane: rozmiar bloku - 2K.B, liczba rekordów pliku:- 1.000.000, rozmiar rekordu pliku:- 256B, rozmiar pliku : 300MB. rozmiar adresu rekordu - 4B. rozmiar indeksowanego atrybutu: 20B.

Zad. 7. Przedstaw przykład realizacji **konfliktowo-uszeregowalnej** $r(\tau)$. zawierającej operacje należące do co najmniej 3 transakcji T1, T2 i T3, o następujących cechach:

1. każda operacja transakcji T1 poprzedza w realizacji $r(\tau)$ każdą operację transakcji T2, każda operacja transakcji T2 poprzedza w realizacji $r(\tau)$ każdą operację transakcji T3, i
2. w grafie konfliktowej uszeregowalności $CSRG(r(\tau))$ $T3 \rightarrow T2$ i $T2 \rightarrow T1$

Dla podanej realizacji przedstaw konfliktowo równoważnej realizację sekwencyjną.

Pomoc: Aby skonstruować taką realizację konieczne jest uwzględnienie dwóch dodatkowych transakcji T4 i T5.

Zad. 8. Dana jest następująca realizacja transakcji:

S: $r1(a, 1)$ $w1(a, 3)$ $r2(d, 2)$ $r3(b, 4)$ $w2(d, 5)$ $w3(b, 6)$ CKPT

$R3(c, 7)$ $r2(e, 8)$ $w2(e, 9)$ T2:commit $r3(e, 6)$ $w3(c, 11)$ **crash**

- Przedstaw sekwencję rekordów logu, zapisaną do pliku logu, wygenerowaną przez powyższą sekwencję operacji.
- Przedstaw procedury ROLLBACK i ROLL FORWARD zainicjowane przez system w celu

odtworzenia spójnego stanu bazy danych po awarii,
zakładając punkt kontrolny spójny z pamięcią podręczną.

Zad. 9. Dany jest zbiór transakcji $\tau = \{T_1, T_2\}$, gdzie $T_1 = r_1(x), w_1(x), r_1(y), w_1(z), c_1$ i $T_2 = r_2(y), w_2(y), w_2(z), c_2$. Podaj przykład realizacji:

1. sekwencyjnej obu transakcji
2. współbieżnej nie uszeregowalnej
3. współbieżnej uszeregowalnej ale nie odtwarzalnej
4. współbieżnej uszeregowalnej i odtwarzalnej

gdzie przez **odtwarzalność** rozumiemy, że w przypadku konieczności wycofania jednej z transakcji nie pociąga to za sobą konieczności wycofania drugiej z nich.

Zad. 10. Załóżmy, że bufor stron dyskowych może pomieścić w danej chwili maksymalnie cztery strony i że początkowo bufor jest pusty. Załóżmy, że każda z danych A, B, C, D, E, F zajmuje w całości jedną stronę dyskową. Poniżej podano sekwencję operacji **wykonywanych** na tych danych:

$R_1(A,1) R_2(B,2) W_1(A,3) R_3(C,4) W_2(B,5) C_2 W_3(C,6) R_4(D,7) R_5(E,8) W_5(E,9) R_6(B,5)$

$R_6(A,3) R_3(F,10) W_3(F,11) W_4(D,12)$

a) Podaj pierwszą operację, podczas której zajdzie potrzeba usunięcia strony z bufora w celu wyczytania innej strony.

b) Załóżmy, że algorytmem wymiany stron w buforze jest LRU i że każda strona jest w użyciu tylko na czas trwania operacji odczytu i zapisu tej strony. Jaka strona zostanie usunięta z bufora w momencie **wykonywania** operacji z punktu (a)?

- c) Czy w momencie operacji G2 (commit) strona zmodyfikowana przez transakcję T2 jest zapisywana do bazy danych zakładając strategię No Force?

Zestaw 2

Zad. 1. Opracuj schemat EER dla Polskiego Związku Piłki Nożnej. Wynik analizy mini-świata jest następujący:

PZPN zamierza gromadzić dane ewidencyjne drużyn piłkarskich, stadionów i meczy w rozgrywkach ligowych. Baza danych PZPN ma przechowywać następujące informacje o drużynach: nazwę drużyny, nazwę sponsora, dane osobowe wszystkich trenerów. Każdy z trenerów jest opisany imieniem, nazwiskiem, numerem telefonu i funkcją (np. pierwszy trener, drugi trener).

Drużyny składają się z zawodników, o każdym z nich należy przechować następujące dane: imię, nazwisko, pensję, numer na koszulce i preferowaną pozycję na boisku (np. napastnik, obrońca, stoper).

Każdy stadion, który znajdzie się w ewidencji PZPN będzie opisany nazwą, lokalizacją (pełen adres), maksymalną pojemnością wyrażoną liczbą kibiców.

W celu prowadzenia rankingu ligi, PZPN będzie rejestrował każdy rozegrany mecz. Oprócz informacji, które drużyny rozegrały dany mecz i na jakim stadionie, mecz będzie opisany datą i godziną rozgrywki i wynikiem. Dla każdej drużyny rozgrywającej mecz należy dodatkowo zapamiętać jej aktualny skład i pozycje każdego gracza.

Przedstaw diagram EER odpowiadający powyższemu opisowi miniświata. Określ i zaznacz w modelu identyfikatory encji.

Zad. 2. Przetransformuj poniższy diagram encji-związków (EER) do schematu relacyjnego. Zaznacz klucze główne, klucze obce w schematach relacji, oraz, dodatkowo, wskaż atrybuty z wartościami obowiązkowymi i opcjonalnymi.

Zad. 3. Dany jest następujący schemat pojęciowy. Każdy pacjent ma swojego lekarza. Każdy lekarz posiada specjalność.

Schemat bazy danych obejmuje 3 poniższe relacje:

SPECJALNOŚCI

S_SYMBOL varchar2(3) primary key S_NAZWA varchar2(20) not null

LEKARZE

L_ID_LEKARZA number(3,0) primary key

L_NAZWISKO varchar2(20) not null

L_IMIE . varchar2(20) not null

L_DATA_ZATR datę not null

L_PENSJA number(7,2) not null

L_S_SYMBOL varchar2(3) not null references SPECJALNOŚCI(S_SYMBOL)

PACJENCI

P_ID_PACJENTA number(6,0) primary key,

P_NAZWISKO varchar2(20) not null,

P_IMIE varchar2(20) not null,

P_WIEK number(3,0) not null,

P_L_ID_LEKARZA number(3,0) not null references LEKARZE

(L_ID_LEKARZA)

Sformułuj w języku SQL następujące zapytania wyrażone poniżej w języku naturalnym.

a) Podaj imiona i nazwiska trzech najwcześniej zatrudnionych lekarzy o specjalności 'OKULISTA', (przyjmij, że daty zatrudnienia są unikalne). Wyniki uporządkuj rosnąco wg daty zatrudnienia.

b) Dla każdego lekarza, który jest lekarzem prowadzącym co najmniej jednego pacjenta, podaj imię i nazwisko lekarza, nazwę jego specjalności oraz liczbę prowadzonych przez niego pacjentów.

c) Usuń najwięcej zarabiającego lekarza (lekarzy) spośród lekarzy, którzy nie prowadzą żadnego pacjenta.

d) Podnieś o 10% pensję lekarzom, których zarobki mieszczą się w przedziale od 2500 do 2800 (włącznie), a nazwa ich specjalności kończy się na literę 'A'.

Zad. 4. Dana jest relacja R przedstawiona poniżej,

R

A	B	C
a1	b1	c1
a1	b1	c2
a2	b1	c1
a2	b1	c2

Relacja R została zdekomponowana na dwie relacje R1 i R2, takie, że R1(A. B), R2(B. C). R1

A	B
a1	b1
a2	b1

R2

B	C
b1	c2
b1	c1

1. Czy połączenie naturalne relacji R1 i R2 daje w wyniku oryginalną relację R?
 2. Czy dekompozycja relacji R na relacje R1 i R2 jest dekompozycja bez utraty informacji?
- Uzasadnij odpowiedź.

Zad. 5. Dany jest schemat relacji R =(A, B, C, D, E, F, G, H, I). Dany jest również następujący zbiór zależności funkcyjnych dla schematu R:

Podaj klucz (klucze) schematu R.

Schemat relacji R jest w pierwszej postaci normalnej. Zdekomponuj schemat R do zbioru schematów u 2NF. ale nie 3NF. Uzasadnij, że otrzymane schematy są w 2NF i że nie są w 3NF. Zdekomponuj otrzymany zbiór schematów w kroku 2 do zbioru schematów w 3NF. Czy otrzymane w wyniku dekompozycji schematy są również w BCNF?

AB → CD

A → E B → F H C → G

D → B

G → C

H → I

Zad. 6. Na nieuporządkowanym pliku na nieunikalnym atrybucie o selektywności 1% założono indeks B+ drzewo. Oblicz, zakładając równomierny rozkład wartości indeksowanego atrybutu, ile w najgorszym przypadku należy wykonać operacji I/O w celu odczytania, z **wykorzystaniem** indeksu, wszystkich rekordów pliku z zadaną wartością indeksowanego atrybutu.

Dane: rozmiar bloku - 2KB, liczba rekordów pliku:- 1.000.000, rozmiar rekordu pliku:- 256B. rozmiar pliku : 300MB. rozmiar adresu rekordu - 4B, rozmiar indeksowanego atrybutu: 20B.

Zad. 7. Przedstaw przykład realizacji **konfliktowo-uszeregowalnej** $r(\tau)$, zawierającej operacje należące do co najmniej 3 transakcji T1, T2 i T3, o następujących cechach:

1. każda operacja transakcji T1 poprzedza w realizacji $r(\tau)$ każdą operację transakcji T2, każda operacja transakcji T2 poprzedza w realizacji $r(\tau)$ każdą operację transakcji T3, i 2. w grafie konfliktowej uszeregowalności CSRG($r(\tau)$) $T3 \rightarrow T2$ i $T2 \rightarrow T1$

Dla podanej realizacji przedstaw konfliktowo równoważną jej realizację sekwencyjną.

Pomoc: Aby skonstruować taką realizację konieczne jest uwzględnienie dwóch dodatkowych transakcji T4 i T5.

Zad. 8. Dana jest następująca realizacja transakcji:

S: $r_2(a, 1)$ $r_1(b, 1)$ $w_1(b, 3)$ $r_3(c, 4)$ $w_2(a, 5)$ $w_3(c, 6)$ CKPT $r_4(d, 7)$ $r_2(e, 8)$ $w_2(e, 9)$ T2:commit $r_4(e, 9)$ $w_4(d, 11)$ **crash**

- Przedstaw sekwencję rekordów logu, zapisaną do pliku logu, wygenerowaną przez powyższą sekwencję operacji.
- Przedstaw procedury ROLLBACK i ROLL FORWARD zainicjowane przez system w celu odtworzenia spójnego stanu bazy danych po awarii, zakładając punkt kontrolny spójny z pamięcią podręczną.

Zad. 9. Dany jest zbiór transakcji $\tau = (T1, T2)$, gdzie $T1 = r_1(x), w_1(x), r_1(y), w_1(z), c1$ i $T2 = r_2(y), w_2(y), w_2(z), c2$. Podaj przykład realizacji:

1. **sekwencyjnej** obu transakcji
2. **współbieżnej** nie uszeregowalnej
3. **współbieżnej** uszeregowalnej ale nie odtwarzalnej
4. **współbieżnej** uszeregowalnej i odtwarzalnej

gdzie przez **odtwarzalność** rozumiemy, że w przypadku konieczności wycofania jednej z transakcji nie pociąga to za sobą konieczności wycofania drugiej z nich.

Zad. 10. Załóżmy, że bufor stron dyskowych może pomieścić w danej chwili maksymalnie cztery strony i że początkowo bufor jest pusty. Załóżmy, że każda z danych A, B, C, D, E, F zajmuje w całości jedną stronę dyskową. Poniżej podano sekwencję operacji **wykonywanych** na tych danych: $R_1(A,1)$ $R_2(B,2)$ $W_1(A,3)$ $R_3(C,4)$ $W_2(B,5)$ C_2 $W_3(C,6)$ $R_4(D,7)$ $R_5(E,8)$ $W_5(E,9)$ $R_6(B,5)$ $R_3(F,10)$ $W_3(F,11)$ $W_4(D,12)$

- a) Podaj pierwszą operację, podczas której zajdzie potrzeba usunięcia strony z bufora w celu wczytania innej strony.
- b) Załóżmy, że **algorytmem** wymiany stron w buforze jest LRU i że każda strona jest w użyciu tylko na czas trwania operacji odczytu i zapisu tej strony. Jaka strona zostanie usunięta z bufora w momencie wykonywania operacji z punktu (a)?
- c) Czy w momencie operacji Q (commit) strona zmodyfikowana przez transakcję π jest zapisywana do bazy danych zakładając strategię No Force?