

**Zad. 1.** Narysuj schemat EER dla przedstawionej poniżej rzeczywistości. Oznacz unikalne identyfikatory encji oraz opcjonalność/obowiązkowość atrybutów. Dla każdego związku zaznacz jego opcjonalność/obowiązkowość oraz stopień.

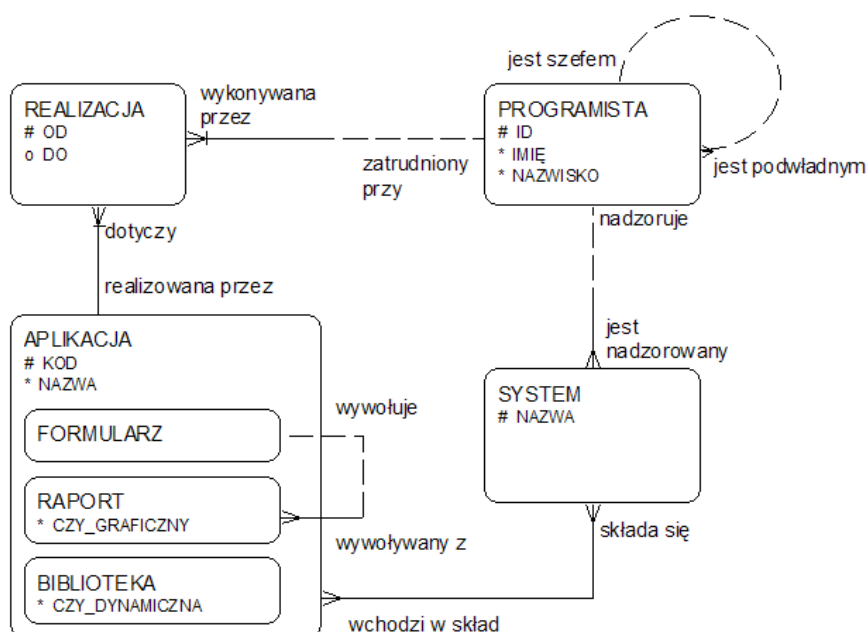
Biuro podróży „FLY & RETURN ALIVE” specjalizuje się w organizowaniu wakacji zagranicą. Oferta podzielona jest na pakiety wakacyjne, każdy pakiet posiada unikalną nazwę (np. „Grecja’2004”) i kategorię (np. „wczasy rodzinne”, „sporty ekstremalne”, itd.). Biuro utrzymuje katalog hoteli. Każdy hotel posiada nazwę, kategorię, dane o lokalizacji (kraj i miasto) i krótki opis najciekawszych miejsc w okolicy. Każdy pokój w hotelu posiada numer (unikalny w obrębie danego hotelu) oraz typ („pokój pojedynczy”, „pokój podwójny”, „apartament”).

Każdy pakiet wakacyjny ma przypisaną pulę pokoi w poszczególnych hotelach w określonych okresach czasu. Jeden pokój może być przypisany kilkakrotnie do pakietów wakacyjnych w rozłącznych okresach czasu. Cena za ten sam pokój może być różna (zależy od pakietu, do którego pokój jest przypisany, oraz od okresu pobytu). Np. cena za pojedynczy pokój nr 90 w hotelu Posejdon w Grecji w terminie od 15 - 30 lipca to 1 500 zł, natomiast w terminie 1 - 15 września cena spada do 1 250 zł.

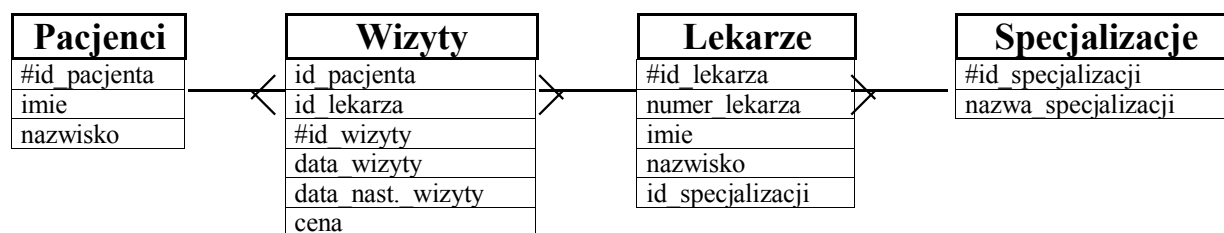
Każdy klient biura podróży posiada unikalny numer PESEL, imię, nazwisko oraz adres, składający się z nazwy ulicy, numeru domu, numeru mieszkania, nazwy miasta i kodu pocztowego. Dla każdego klienta przechowywana jest również informacja o procentowej zniżce przy zakupie oferty. Klienci dokonują rezerwacji pobytów w pokojach hotelowych, przypisanych do wybranych pakietów wakacyjnych w wybranych okresach czasu. Dokonanie rezerwacji powinno wiązać się z wniesieniem częściowej opłaty za pokój. Reszta ceny powinna zostać wpłacona maksymalnie na 2 tygodnie przed wyjazdem. Wszystkie wpłaty klientów powinny być ewidencjonowane. Dodatkowo, system powinien umożliwiać przechowywanie „rezerwowej” listy klientów dla danego pakietu. Są to klienci, którzy nie zarezerwowali pokoju (nie było już wolnych), ale liczą na wycofanie rezerwacji innych klientów.

**Zad. 2.** Przetransformuj poniższy diagram związków encji (EER) do schematu relacyjnego.

Zaznacz klucze główne (podkreśleniem ciągłym) oraz klucze obce (podkreśleniem przerywanym) w schematach relacji. Zaznacz dodatkowo obowiązkowość/opcjonalność kolumn relacji.



**Zad. 3.** Dany jest następujący schemat bazy danych:



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

- Wyświetl imiona i nazwiska tych pacjentów, którzy w styczniu 2004 roku korzystali z porady okulisty i będą musieli przyjść na wizytę kontrolną.
- Znajdź imiona i nazwiska tych lekarzy, którzy w latach 2000-2003 wykonali sumarycznie najwięcej wizyt (tylko 5 najbardziej zapracowanych).
- Podnieś o 22% ceny tych wszystkich wizyt u lekarzy ortopedów, które miały miejsce w 2003 roku.
- Znajdź tych pacjentów, którzy w 2003 roku nie odwiedzili żadnego lekarza ani nie mieli wyznaczonej żadnej wizyty na 2003 rok.

**Zad. 4.** Dane są trzy relacje  $t(T)$ ,  $t_1(T_1)$ ,  $t_2(T_2)$ , o schematach  $T$ ,  $T_1$  i  $T_2$ , takie, że  $t_1 \text{ Join } t_2 = t$ . Załóżmy ponadto, że  $T_1 \cap T_2 \rightarrow T_2$ , to jest, iloczyn logiczny zbiorów atrybutów schematów  $T_1$  i  $T_2$  wyznacza funkcyjnie schemat  $T_2$ . Załóżmy ponadto, że relacja  $t$  została zdekomponowana na dwie podrelacje  $s_1(S_1)$  i  $s_2(S_2)$ , o schematach  $S_1$  i  $S_2$ , takie, że  $S_1 = T_1$  i  $S_2 = T_2$ . Dekompozycję uzyskano przez wykonanie projekcji na relacji  $t$ :

$$s_1 = \Pi_{S_1}(t) \text{ i } s_2 = \Pi_{S_2}(t),$$

gdzie  $S_1$  i  $S_2$  są schematami relacji  $s_1$  i  $s_2$ .

- Podaj przykład ilustrujący tezę, że relacje  $s_1$  i  $s_2$  nie muszą być identyczne z relacjami  $t_1$  i  $t_2$ ,  $s_1 \neq t_1$  i  $s_2 \neq t_2$ .

Czy jeżeli zastąpimy operację Join operacją Outer Join (obustronny outer join), tj.  $t = t_1 \text{ Outer Join } t_2$  (reszta opisu identyczna), to czy musi zachodzić  $s_1 = t_1$  i  $s_2 = t_2$ . TAK czy NIE. Uzasadnij krótko odpowiedź.

**Zad. 5.** Dana jest tablica **Olimpijczycy** (Rysunek 1). Dany jest następujący zbiór zależności funkcyjnych między atrybutami tablicy **Olimpijczycy**:

symbolKraju  $\rightarrow$  nazwaKraju

symbolDyscypliny  $\rightarrow$  nazwaDyscypliny

idTrenera  $\rightarrow$  nazwiskoTrenera

idZawodnika, symbolKraju  $\rightarrow$  nazwisko, imię, numer, idTrenera, nazwiskoTrenera

symbol Kraju	nazwa Kraju	id Zawodnika	nazwisko	imię	Numer	symbol Dysc.	nazwa Dysc.	id Trenera	nazwisko Trenera
PL	Polska	9001	Niebieski	Jan	4590	SD SP	skok w dal sprint	100	Dobry
GER	Niemcy	9002	Grün	Johann	2342	SP ST	sprint sztafeta	200	Słaby
FR	Francja	9001	Rouge	Marc	2190	T	tenis	100	Wspaniały

Rysunek 1. Tablica **Olimpijczycy**

- Zdekomponuj schemat do 1NF.
- Podaj klucz (klucze) schematu relacji **Olimpijczycy**?

3. Czy schemat relacji **Olimpijczycy** jest w 2NF? Uzasadnij swoją odpowiedź. Jeśli schemat nie jest w 2NF, zdekomponuj schemat do 2NF.
4. Czy schemat relacji **Olimpijczycy** jest w 3NF? Uzasadnij swoją odpowiedź. Jeśli schemat nie jest w 3NF, zdekomponuj schemat do 3NF.
5. Załóżmy, że schemat relacji **Olimpijczycy** został zdekomponowany na dwa podschematy
  - **Kraje**(symbolKraju, nazwaKraju, idTrenera, nazwiskoTrenera)
  - **Zawodnicy**(idZawodnika, symbolKraju, nazwisko, imię, numer, symbolDyscypliny, nazwaDyscypliny).

Czy dekompozycja schematu **Olimpijczycy** na podschematy **Kraje** i **Zawodnicy** jest dekompozycją bez utraty informacji? Uzasadnij odpowiedź.

**Uwaga:** podczas dekompozycji proszę nie przepisywać przykładowych danych, wystarczy nazwa relacji, lista atrybutów i zaznaczone zależności funkcyjne.

**Zad. 6.** Na nieuporządkowanym pliku, na nieunikalnym atrybucie założono indeks B+ drzewo. Oblicz, zakładając równomierny rozkład wartości indeksowanego atrybutu i równomierne rozmieszczenie rekordów w pliku, ile **średnio** 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 w pliku - 1.000.000, rozmiar rekordu pliku - 256B, rozmiar pliku - 300MB, rozmiar adresu rekordu - 4B, rozmiar adresu bloku - 4B, rozmiar indeksowanego atrybutu - 20B, liczebność dziedziny indeksowanego atrybutu – 1.000.

**Zad. 7.** Zarządzanie współbieżnością.

**(A)** 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.

**(B)** Dana jest następująca realizacja:

$r$ : T1:r(z) T3:r(x) T2:r(z) T1:w(z) T1:w(y) T1:c T2:w(y) T2:w(u) T2:c T3:w(y) T3:c

Pokaż, że realizacja  $r \in SR - CSR$ , gdzie SR oznacza zbiór realizacji uszeregowalnych, natomiast CSR oznacza zbiór realizacji konfliktowo-uszeregowalnych (D-uszeregowalnych).

**Zad. 8.** Dana jest następująca realizacja transakcji:

S: T1: r(A, 50) T1:r(D, 10) T1:w(D, 20) T1:c T2:r(B, 10) CKPT Flush  
(D)

T2:w(B, 15) T4:r(D, 20) T4:w(D, 30) T3:w(C, 30) Flush(B) CKPT  
T4:r(A, 50) T4:w(A, 20) T4:c Flush(D) T2:r(D, 30) T2:w(D, 25)

**awaria**

Operacja **Flush**(strona) powoduje zapis danej strony na dysk (załóżmy, że operacja ta nie generuje rekordu logu).

- 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,
- Podaj stan bazy danych po zakończeniu procesu odtwarzania:  
 $X = ? \quad Y = ? \quad Z = ? \quad P = ?$

zakładając punkt kontrolny spójny z pamięcią podręczną.

**Zad. 9.**

**(A)** Dane są dwie relacje R i S o rozmiarach odpowiednio, R – 1000 bloków, S – 500 bloków. Załóżmy, że rozmiar bufora danych dostępny dla operacji połączenia wynosi 101 bloków.

- Podaj koszt wykonania operacji połączenia metodą block nested-loop liczoną w liczbie operacji I/O. Wyjaśnij w jaki sposób został obliczony koszt.
- Podaj koszt wykonania operacji połączenia metodą hash-join liczoną w liczbie operacji I/O. Wyjaśnij w jaki sposób został obliczony koszt.

**(B)** Dane są dwie relacje R i S o rozmiarach odpowiednio, R – 10 000 bloków, S – 10 000 bloków. Podaj minimalny rozmiar bufora danych M, który pozwala na wykonanie operacji połączenia relacji R i S metodą block nested-loop kosztem co najwyżej 25 000 operacji I/O.