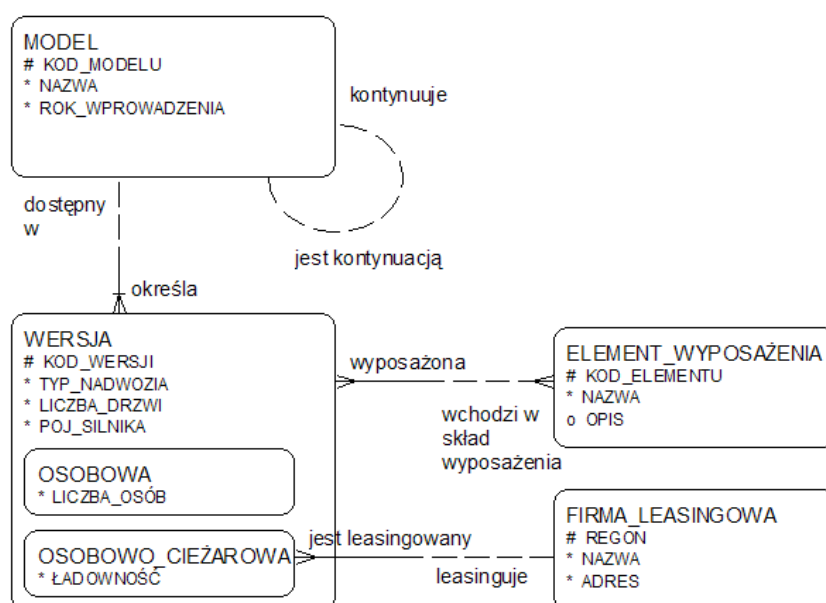


**Zad. 1.** Narysuj schemat EER dla przedstawionej poniżej rzeczywistości. Oznacz unikalne identyfikatory encji oraz opcjonalność/obowiązkowość atrybutów. Dla każdego związku zaznacz jego opcjonalność/obowiązkowość oraz stopień.

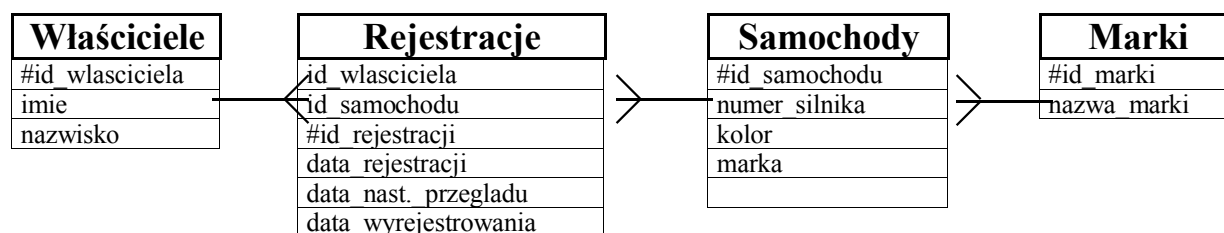
Przedsiębiorstwo transportowe „JELCZ RULES” zajmuje się transportem towarów z hurtowni do sklepów. Przedsiębiorstwo jest odpowiedzialne za odbiór partii towarów z hurtowni i rozwiezienie ich po wskazanych sklepach. System informatyczny przedsiębiorstwa powinien przechowywać następujące informacje:

- Przedsiębiorstwo aktualnie obsługuje 6 hurtowni (liczba ta może ulec zmianie). Każda hurtownia identyfikowana jest przez swoją nazwę, pozostałe informacje to adres hurtowni i numer telefonu.
- Przedsiębiorstwo rozwozi towary do 156 sklepów. Każdy sklep jest identyfikowany przez unikalny identyfikator numeryczny, posiada nazwę, adres i numer telefonu.
- Przedsiębiorstwo wykorzystuje w pracy 245 ciężarówek. Każdy samochód posiada unikalny numer identyfikacyjny, pojemność (wyrażoną w m<sup>3</sup>) oraz ładowność (wyrażoną w tonach).
- W przedsiębiorstwie pracuje 320 kierowców. Każdy kierowca identyfikowany jest przez unikalny numer, posiada imię i nazwisko, adres i telefon kontaktowy.
- Ciężarówki rozwożą towar po zdefiniowanych wcześniej trasach. Podczas jednej trasy określona ciężarówka, prowadzona przez wskazanego kierowcę, pobiera partie towarów ze wskazanych hurtowni (jedna partia towaru musi zostać pobrana z jednej hurtowni) i dostarcza je do określonych sklepów (towar z jednej partii może trafić do wielu sklepów). Część partii towarów, która trafia do jednego sklepu, nosi nazwę dostawy. Kierowca ciężarówki posiada informacje o wszystkich dostawach w ramach wszystkich partii dla realizowanej trasy. Dla każdej trasy ewidencjonowana jest data realizacji oraz długość (w kilometrach). W ciągu jednego dnia ciężarówka może wykonać kilka tras.

**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 właścicieli sam marki FIAT, którzy lutym 2004 powinni udać się na przegląd techniczny.
- Znajdź imiona i nazwiska tych właścicieli, którzy w latach 1995-1999 zarejestrowali najwięcej samochodów (tylko 5 „najbogatszych” właścicieli).
- Wszystkim samochodom marki TRABANT zarejestrowanym przed 1990 rokiem (i ciągle zarejestrowanym) ustaw datę następnego przeglądu na 1 kwietnia 2004.
- Znajdź te marki samochodów, dla których nie ma już żadnych „aktywnych” rejestracji.

**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 *Studenci* (Rysunek 1). Dany jest następujący zbiór zależności funkcyjnych między atrybutami tablicy *Studenci*

symbolWydziału  $\rightarrow$  nazwaWydziału

symbolJO  $\rightarrow$  nazwaJO

symbolKierunku  $\rightarrow$  nazwaKierunku

idStudenta, symbolWydziału  $\rightarrow$  nazwisko, imię, średnia, symbolKierunku, nazwaKierunku

symbol Wydziału	nazwa Wydziału	id Studenta	nazwisko	imię	średnia	symbol JO	nazwa JO	symbol Kierunku	nazwa Kierunku
WE	Elektryczny	1001	Niebieski	Jan	4.0	A N	angielski niemiecki	A	Automatyka
WIIZ	Informatyka i Zarządzanie	1002	Zielony	Adam	3.6	A R	angielski rosyjski	I	Informatyka
WIIZ	Informatyka i Zarządzanie	1001	Czerwony	Marek	3.8	N H	niemiecki rosyjski	A	Automatyka

Rysunek 1. Tablica *Studenci*

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

3. Czy schemat relacji **Studenci** jest w 2NF? Uzasadnij swoją odpowiedź. Jeśli schemat nie jest w 2NF, zdekomponuj schemat do 2NF.
4. Czy schemat relacji **Studenci** jest w 3NF? Uzasadnij swoją odpowiedź. Jeśli schemat nie jest w 3NF, zdekomponuj schemat do 3NF.
5. Załóżmy, że schemat relacji **Studenci** został zdekomponowany na dwa podschematy
  - **Wydziały**(symbolWydziału, nazwaWydziału, symbolKierunku, nazwaKierunku)
  - **Studenci2**(idStudenta, symbolWydziału, nazwisko, imię, średnia, symbolJO, nazwaJO).
 Czy dekompozycja schematu **Studenci** na podschematy **Wydziały** i **Studenci2** 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 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 i równomierne rozmieszczenie rekordów w pliku, ile w **najgorszym przypadku** 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 - 2KB, liczba rekordów pliku:- 1.000.000, rozmiar rekordu pliku:- 256B, rozmiar pliku : 300MB, rozmiar adresu rekordu - 4B, rozmiar adresu bloku - 4B, rozmiar indeksowanego atrybutu: 20B.

**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 = ? Y = ? Z = ? 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.