

# *Bazy danych*



**Andrzej Łachwa, UJ, 2015**

[andrzej.lachwa@uj.edu.pl](mailto:andrzej.lachwa@uj.edu.pl)

**3/14**

## Specyfikacja wymagań

Zanim rozpoczniemy modelowanie, musimy dokładnie określić obszar analizy oraz zrozumieć go! W praktyce analitycy udają się do przedsiębiorstwa czy instytucji, dla której ma być wykonana baza danych, i tam poprzez wywiady, ankiety i obserwacje poznają specyfikę działania pracowników, obieg dokumentów, procedury, czynności, normy etc. Następnie analitycy dokonują uproszczenia: zastępują skomplikowane i szczegółowe okoliczności występujące w świecie rzeczywistym zrozumiałym modelem danych. Modelowaniu danych towarzyszy zawsze modelowanie struktury organizacyjnej i sieci procesów. Baza danych jest bowiem projektowana dla konkretnych użytkowników i dla konkretnego sposobu korzystania z danych.

Pomocą w zrozumieniu obszaru analizy i poprawnym zaprojektowaniu bazy jest specyfikacja wymagań. Przyjmuje ona (w naszym podejściu) postać tekstu opisującego dane, ich przetwarzanie i użytkowników. W specyfikacji tej powinny się znaleźć także szacunki dotyczące wielkości bazy, liczby użytkowników, intensywności korzystania z danych, używanego sprzętu i oprogramowania. Specyfikacja wymagań stanowi punkt odniesienia całego projektu. Jest ona ważna także dlatego, że w trakcie wykonywania projektu zmienić może się zarówno obszar analizy, jak i wymagania przyszłych użytkowników.

### Przykład: MULTIKINO (fragment specyfikacji)

*W multikinie jest kilka sal. W każdej z nich widownia dzieli się na rzędy i miejsca w rzędzie. Rzędy mogą być różnej długości. Liczba rzędów w różnych salach może być różna.*

*W każdej sali każdego dnia może być kilka seansów. Film może być wyświetlany na wielu seansach w jednej sali, a także w różnych salach równocześnie.*

*Klient może zarezerwować bilety na seans przez Internet. Zarezerwowane bilety wykupuje się w kasie na hasło.*

*Nie zbiera się żadnych informacji o klientach i nie prowadzi się programów lojalnościowych.*

*Dla każdego seansu muszą być ustalone dwie ceny – bilet zwykły i ulgowy. Wszystkie bilety na dany seans dla klientów indywidualnych są sprzedawane w tych cenach. Ceny biletów na różne seanse mogą być różne (ustala się je w zależności od dnia tygodnia, godziny seansu, sali, filmu, pogodny, pory roku etc.).*

*Ceny biletów dla grup zorganizowanych są negocjowane. Miejsca są rezerwowane i sprzedawane w tzw. cenie specjalnej. Na dane seans dla różnych grup mogą być ustalone różne ceny specjalne.*

*Projektowana aplikacja ma wspomagać pracę w kasach multikina, pracę menadżera (ustalanie programów, cen biletów, obsługa grup zorganizowanych, analizy ekonomiczne) oraz umożliwiać rezerwacje biletów przez klientów.*

## Diagram związków encji (ang. *ERD*)

Encja to coś:

- co istnieje niezależnie i jest jednoznacznie identyfikowane, rozpoznawalne,
- o czym informacje mają być przechowywane w bazie,
- co występuje (potencjalnie) w wielu egzemplarzach,
- czemu łatwo nadać nazwę pospolitą (a nie własną)!

Typ encji to „pojemnik” na encje, które są podobne do siebie w tym sensie, że ich charakterystyki (czyli to, co ma być pamiętane w bazie) mają taką samą strukturę.

Umawiamy się, że typom encji nadajemy nazwy pospolite w formie liczby pojedynczej.

Encją może być:

– obiekt fizyczny: ten budynek, książka leżąca na półce, samochód w garażu, pracownik, uczeń, klient banku, Twój pies, obraz, który kupujesz w galerii ...

(np. książki w bibliotece to obiekty fizyczne, konkretne przedmioty; każdy opatrzony unikatowym numerem)

– obiekt abstrakcyjny: książka na witrynie księgarni internetowej, samochód w katalogu, słowo języka polskiego, utwór muzyczny, rachunek bankowy...

(np. książka w księgarni, to obiekt abstrakcyjny, to tytuł reprezentowany wieloma egzemplarzami stojącymi na półce, a także taki, którego aktualnie brak na półce)

– zdarzenie: wypadek komunikacyjny, lekcja w szkole, zakup koszyka towarów w sklepie internetowym, przelew bankowy, seans w kinie, wizyta u lekarza ...

– pojęcie: kolor, styl w architekturze ...

– miejsce: miejsce zamieszkania, miejscowość, obszar geograficzny lub administracyjny ...

– kształt

– proces, czynność

...



Dla przykładowego obszaru analizy (multikino) możemy zaproponować następujące typy encji:

***FILM***

***SEANS***

***SALA***

***MIEJSCE***

***REZERWACJA***

***BILET***

***GRUPA***

Czy na pewno są to typy encji?

Czy potrzebne są inne typy encji w tym obszarze?

Między encjami mogą zachodzić rozmaite związki.

Związki te mogą łączyć encje w pary, trójki, czwórki itd. Mówimy wówczas o związkach stopnia 2-go, 3-go czy 4-tego.

Związki mogą dotyczyć encji tego samego typu (taki jest np. związek typu **PRACOWNIK podlega PRACOWNIK**). Związki takie nazywamy unarnymi.

Związki mogą dotyczyć encji dwóch różnych typów (np. **KLIENT kupuje BILET**); nazywamy je binarnymi.

Przykładem związku łączącego encje trzech typów jest związek opisujący sytuację, gdy lekarz wypisuje pacjentowi receptę. Możemy to zdarzenie modelować jako związek:

***wypisuje (LEKARZ, PACJENT, RECEPTA).***

Jest to związek 3-go stopnia między trzema różnymi typami encji.

Z inną sytuacją mamy do czynienia w przypadku związku modelującego sprzedaż nieruchomości:

***sprzedaje (PODMIOT, PODMIOT, NIERUCHOMOŚĆ).***

Jest to również związek 3-go stopnia, ale zachodzący między dwoma typami encji. W tym przypadku muszą być nazwane role obu podmiotów wchodzących w ten związek (kto sprzedaje i komu sprzedaje).

Związki oznaczamy czasownikami (np. *podlega, kupuje, wydaje*) lub innymi wyrażeniami predykatywnymi (np. *należy do, jest częścią*) w taki sposób, by było to jak najbardziej zrozumiałe, intuicyjne i zarazem proste.

W omawianym obszarze analizy (multikino) możemy zaproponować m.in. wymienione niżej typy związków:

***wyświetla (SEANS, FILM)***

***odbywa się (SEANS, SALA)***

***zawiera (SALA, MIEJSCE)***

***łączy (BILET, MIEJSCE, SEANS)***

***dotyczy (REZERWACJA, MIEJSCE)***

***obejmuje (REZERWACJA, SEANS)***

## Krotności związków unarnych i binarnych

Związki 2-go stopnia dzielimy na trzy rodzaje oznaczane symbolami „1:1”, „1:N” (czytaj: jeden do wielu) i „N:M” (czytaj: wiele do wielu).

Przykładem związku „1:1” jest związek między ważnym w danej chwili dowodem rejestracyjnym a pojazdem (jeden pojazd może mieć tylko jeden ważny dowód rejestracyjny i na odwrót: ważny dowód rejestracyjny przypisany jest dokładnie do jednego pojazdu).

Związek między dowodem rejestracyjnym a pojazdem jest natomiast związkiem typu „N:1”, bo dla jednego pojazdu wydaje się kolejno, wiele dowodów rejestracyjnych.

## Udział encji w związku

Udział encji danego typu w związku encji może być udziałem pełnym lub częściowym. Udział pełny oznacza, że każda encja danego typu musi wchodzić w taki związek. Udział częściowy oznacza dowolność: niektóre encje wchodzą w ten związek, inne nie.

W omawianym wyżej drugim związku dowodów rejestracyjnych i pojazdów udział encji typu **DOWÓD** jest pełny, bo każdy dowód jest wydawany dla jakiegoś pojazdu. Udział encji typu **POJAZD** jest częściowy, bo pojazd, przed pierwszym zarejestrowaniem nie posiada żadnego dowodu rejestracyjnego i dowolnym momencie może być wycofany z ruchu (wyrejestrowany).

## Warunek strukturalny udziału encji w związku

Bardziej ogólną charakterystykę związku encji (od omówionych wyżej określeń liczebności i udziału) dają tzw. warunki strukturalne. Warunek taki ma postać  $(k,n)$  i oznacza, że co najmniej  $k$  encji i co najwyżej  $n$  encji danego typu wchodzi w ten związek.

Przykładem warunku strukturalnego jest warunek  $(2,3)$  udziału encji typu *doktorant* w związku z encjami typu *recenzent*. Oznacza on, że każdy doktorant musi mieć dwie recenzje swojej rozprawy, ale zdarza się, że ma 3 recenzje. Większej liczby recenzji nie dopuszcza się.

## Atrybuty

Wszystkie encje danego typu mają charakterystyki o tej samej strukturze: są to wybrane przez analityka istotne własności encji tego typu, zwane dalej atrybutami, tworzące strukturę drzewa, którego korzeniem jest dana encja. Charakterystykę konkretnej encji tworzy układ wartości atrybutów.

Związki encji mogą również posiadać tego rodzaju charakterystyki: układ wartości atrybutów.

Atrybuty typów encji i związków encji mogą być atrybutami złożonymi bądź elementarnymi (liście drzewa). Dla każdego atrybutu elementarnego trzeba określić dziedzinę wartości tego atrybutu.



Atrybuty mogą być jednowartościowe albo wielowartościowe. Ponadto wyróżnia się atrybuty wyliczeniowe (ich wartości są wyliczane z wartości innych atrybutów).

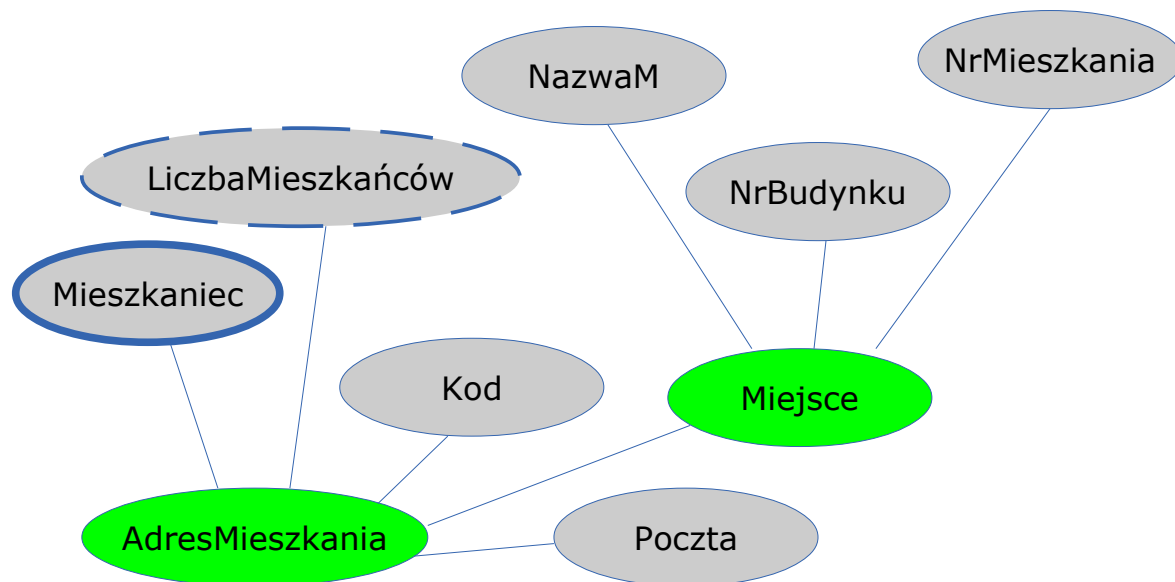
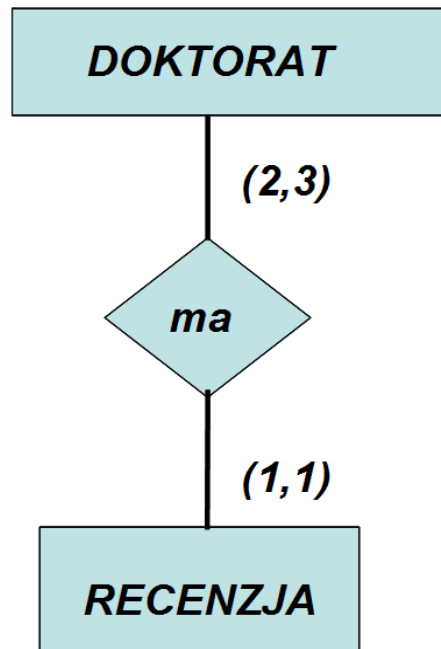
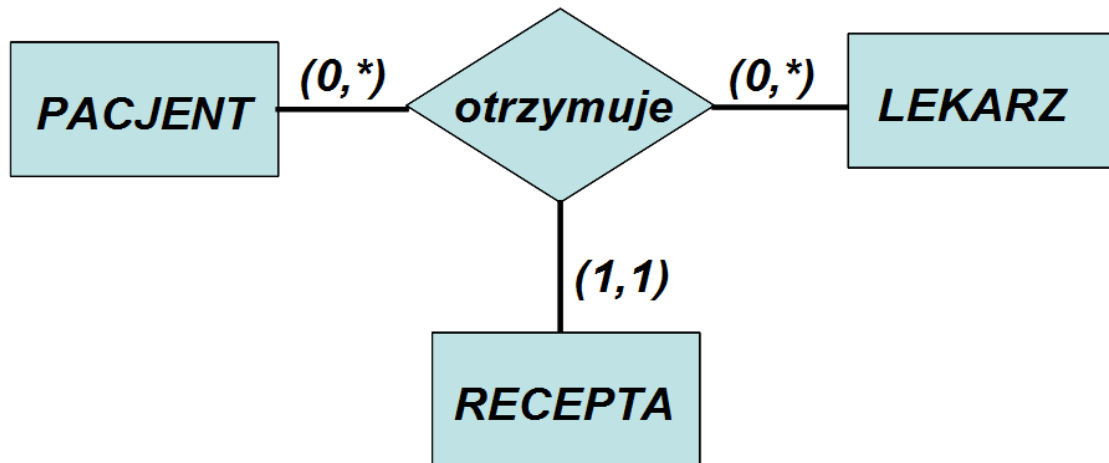


Diagram związków encji jest grafem reprezentującym strukturę danych. Wierzchołki tego grafu reprezentują typy encji, związki i atrybuty, a dla ich odróżniania używa się odpowiednio kształtów prostokąta, rombu i elipsy. Krawędzie jednego typu wiążą prostokąty lub romby z owalami. Krawędzie drugiego typu wiążą prostokąty z rombami. Niektóre z krawędzi drugiego typu mogą być etykietowane.





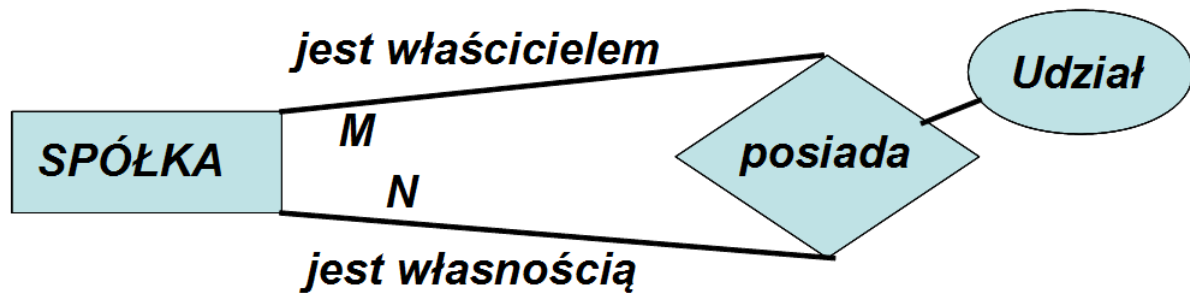


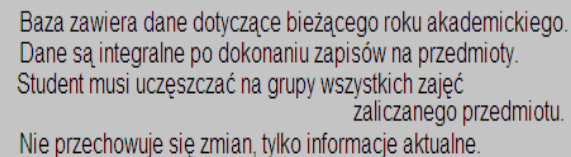
Diagram związków encji (*Entity Relationship Diagram*) to przyjęty w środowisku projektantów baz danych sposób opisywania świata rzeczywistego, to zbiór koncepcji używanych do opisanie struktury danych składających się na projektowaną bazę.

Prezentowana notacja elementów diagramu *ERD* została wprowadzona przez Petera Chena (1976) i jest bardzo popularna.

W wielu narzędziach do rysowania diagramów *ERD* mamy do wyboru także inne notacje (które warto poznać).

Na zajęciach będziemy używać edytora *yEd* firmy *yWorks* (<http://www.yworks.com/en/products/yfiles/yed/>).

Andrzej Łachwa 2007



# Studium przypadku: komunikacja autobusowa

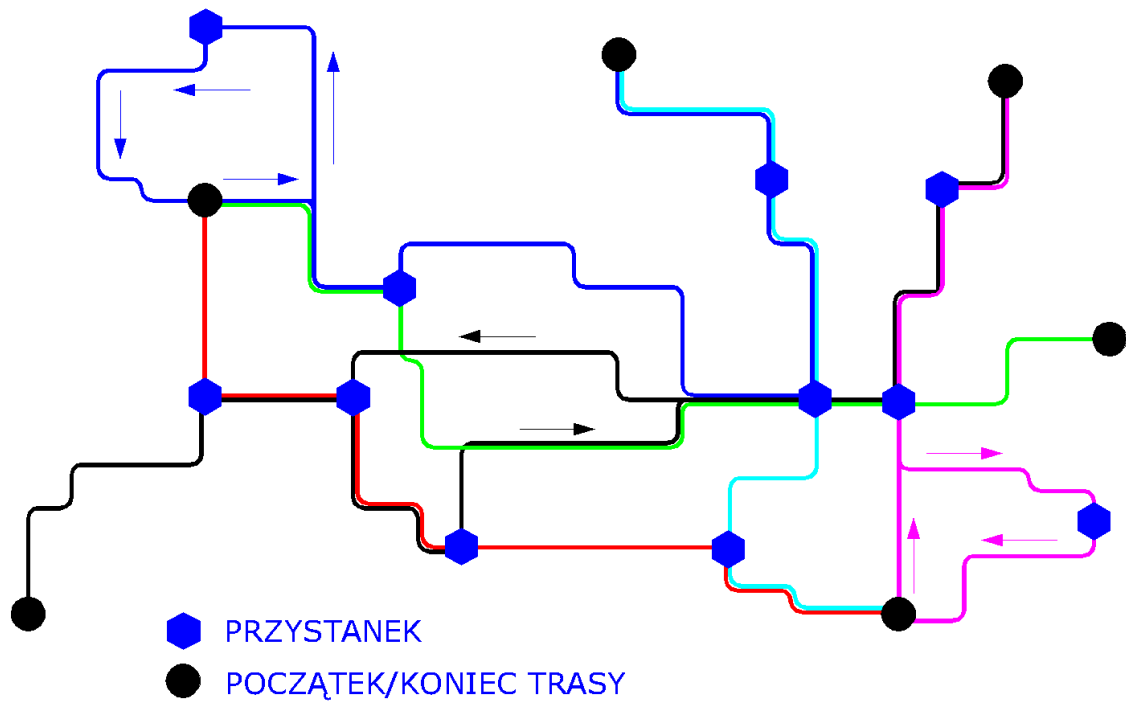
obszar analizy, specyfikacja funkcjonalna i techniczna

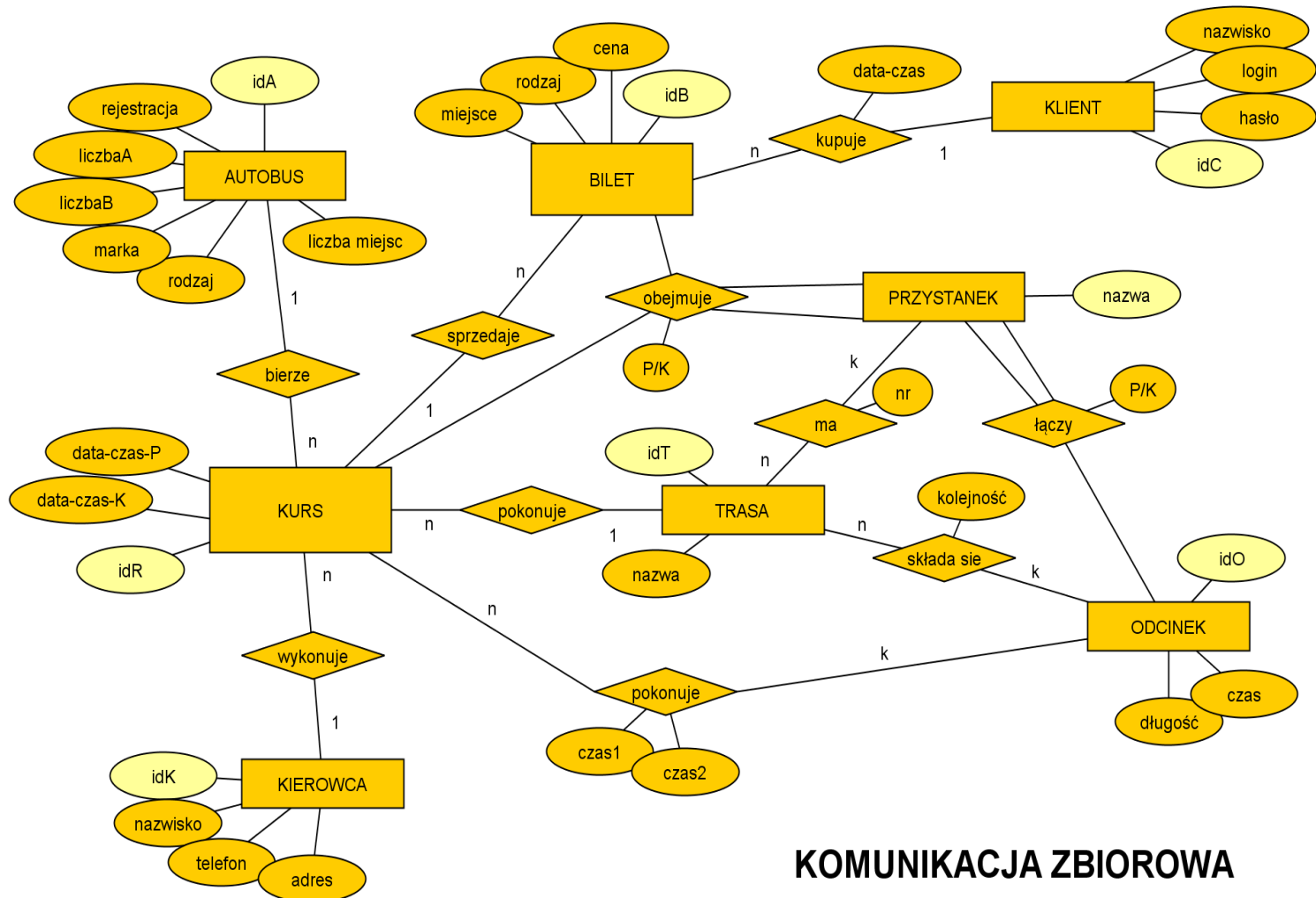
projekt koncepcyjny (diagram związków encji)

projekt logiczny (schemat bazy danych)

projekt fizyczny (skrypt definiujący w dialekcie MySQL)

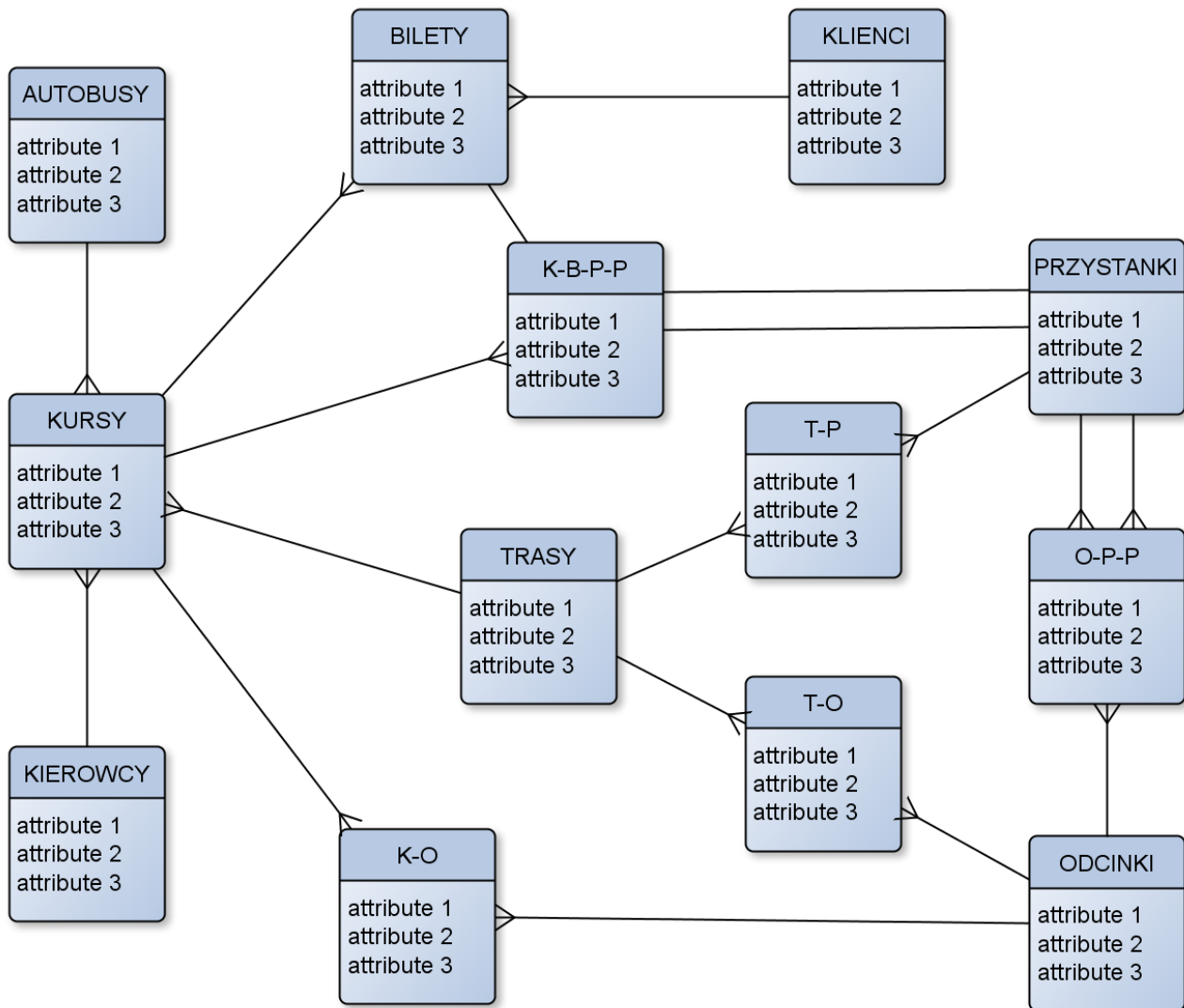






# KOMUNIKACJA ZBIOROWA

Andrzej Łachwa, 2013



## SKRYPT DEFINIUJACY BAZĘ (fragment)

```
CREATE TABLE Kurs (  
    KodKur SMALLINT AUTO_INCREMENT,  
    KodTr VARCHAR(10),  
    Dzień DATE,  
    Godzina TIME,  
    KierID SMALLINT,  
    AutoID SMALLINT,  
    PRIMARY KEY (KodKur)  
);
```

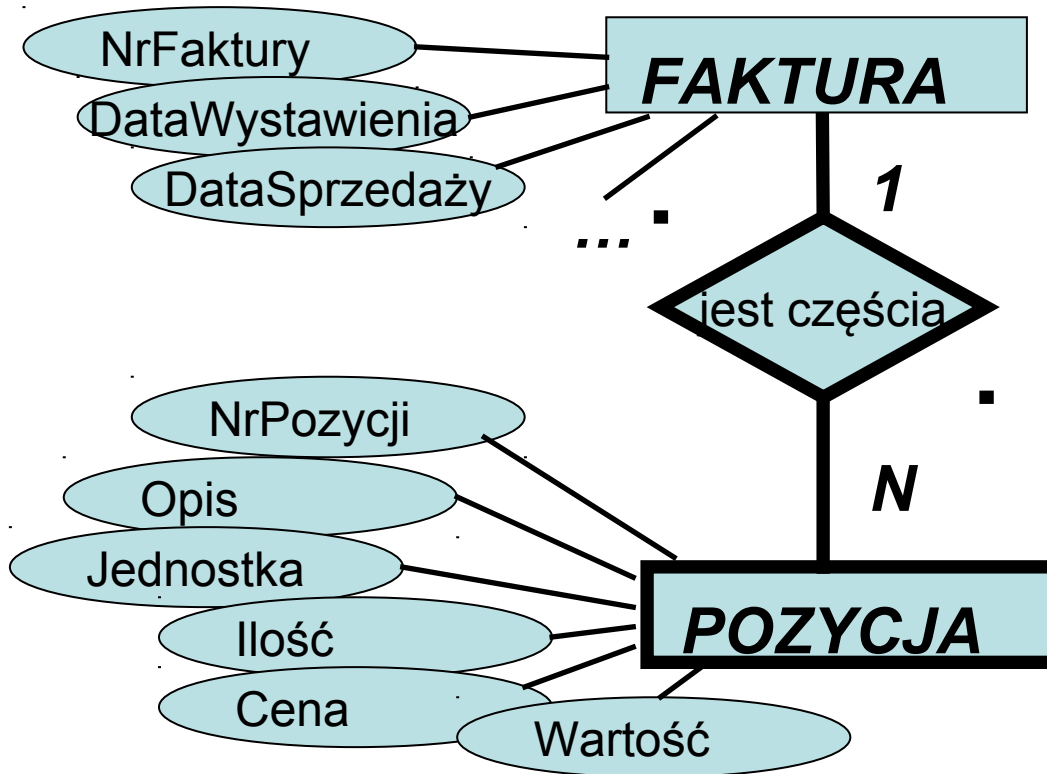
## SKRYPT WYPEŁNIAJĄCY BAZĘ (fragment)

```
INSERT INTO Kurs (KodTr, Dzień, Godzina, KierID, AutoID) VALUES  
('KR-WA',DATE_ADD('2012-07-01', INTERVAL  
    FLOOR(250*RAND()) DAY),    MAKETIME(24*RAND(),  
    60*RAND(), 00), FLOOR(10*RAND())+1,  
    FLOOR(10*RAND())+1),  
('WA-KR',DATE_ADD('2012-07-01', INTERVAL  
    FLOOR(250*RAND()) DAY),    MAKETIME(24*RAND(),  
    60*RAND(), 00), FLOOR(10*RAND())+1, FLOOR(10*RAND())  
    +1),
```

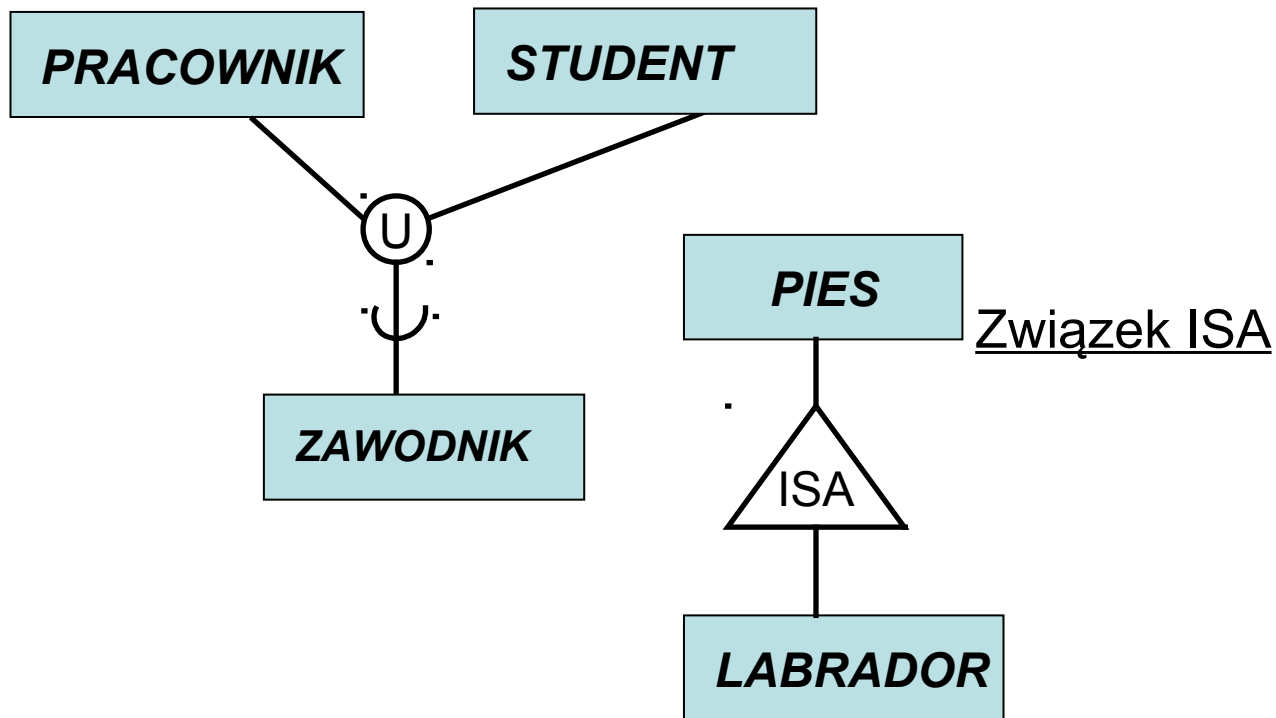
## Rozszerzenia ERD (EER)

- słabe typy encji i związki identyfikujące
- specjalizacja i podklasy
- związki „isa”
- kategoryzacja

## Słaby typ encji i związek identyfikujący

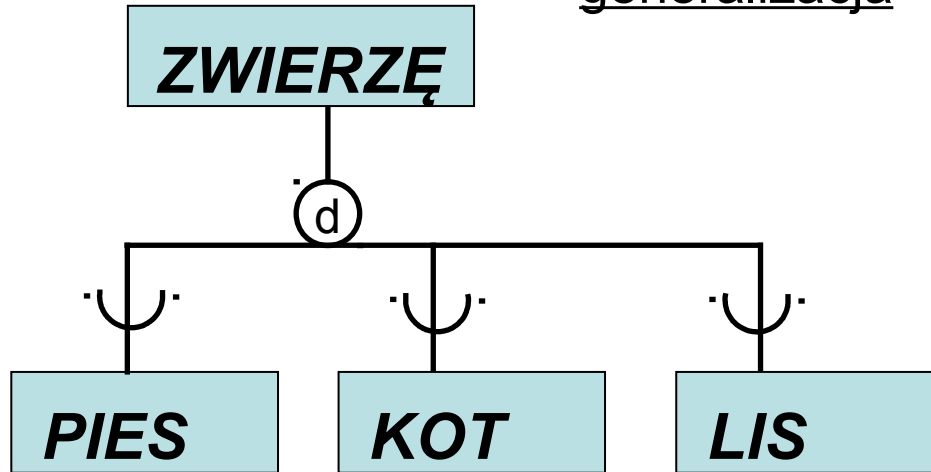


## Kategoryzacja





Specjalizacja /  
generalizacja



specjalizacja nakładająca



specjalizacja rozłączna

