



Bazy Danych

Modele danych

Krzysztof Regulski

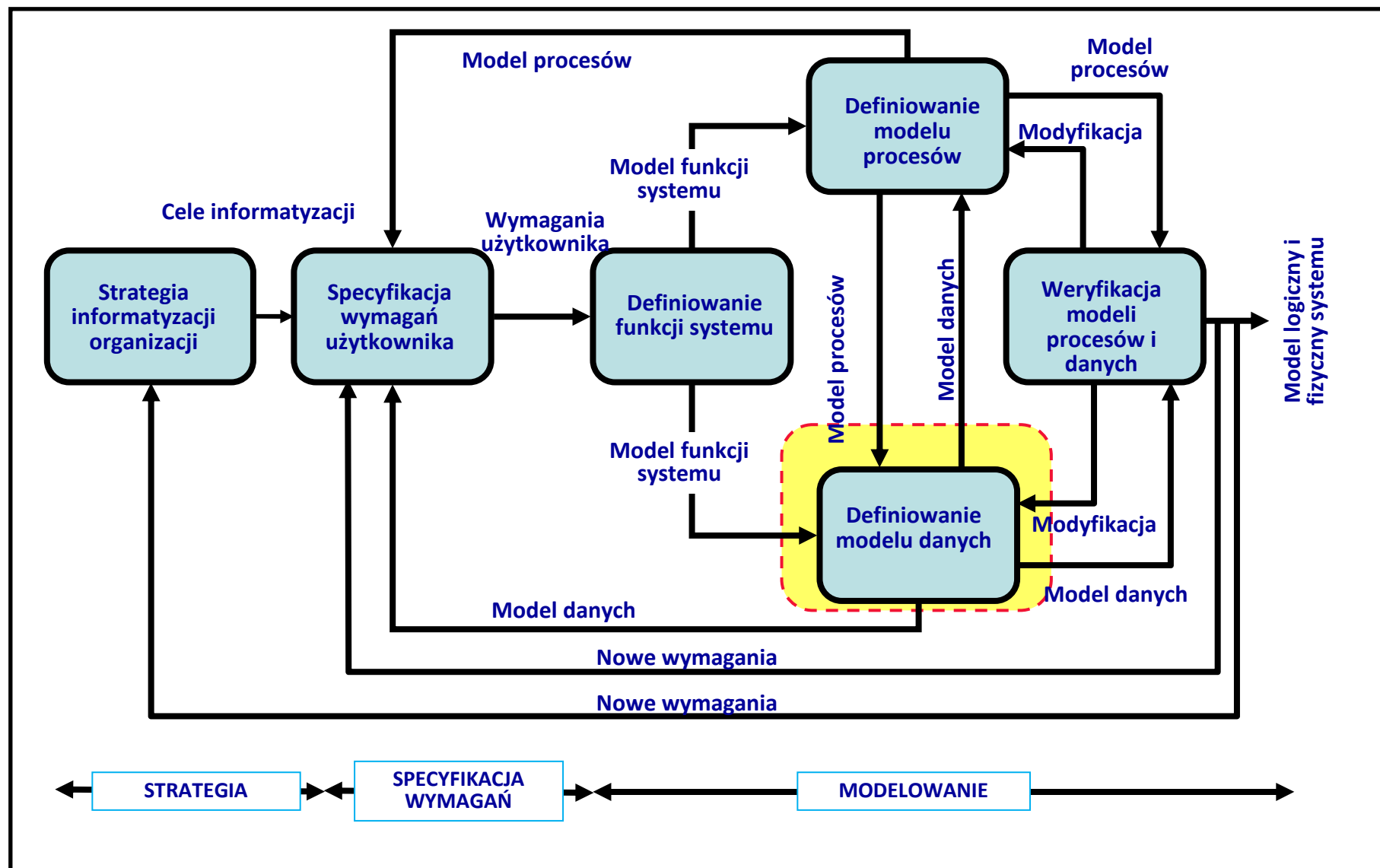
WIMiP, KISiM,

regulski@metal.agh.edu.pl

Konsultacje: wtorek, godz. 15:00 – 16:30

B5, pok. 409

Cele modelowania



Cele modelowania danych

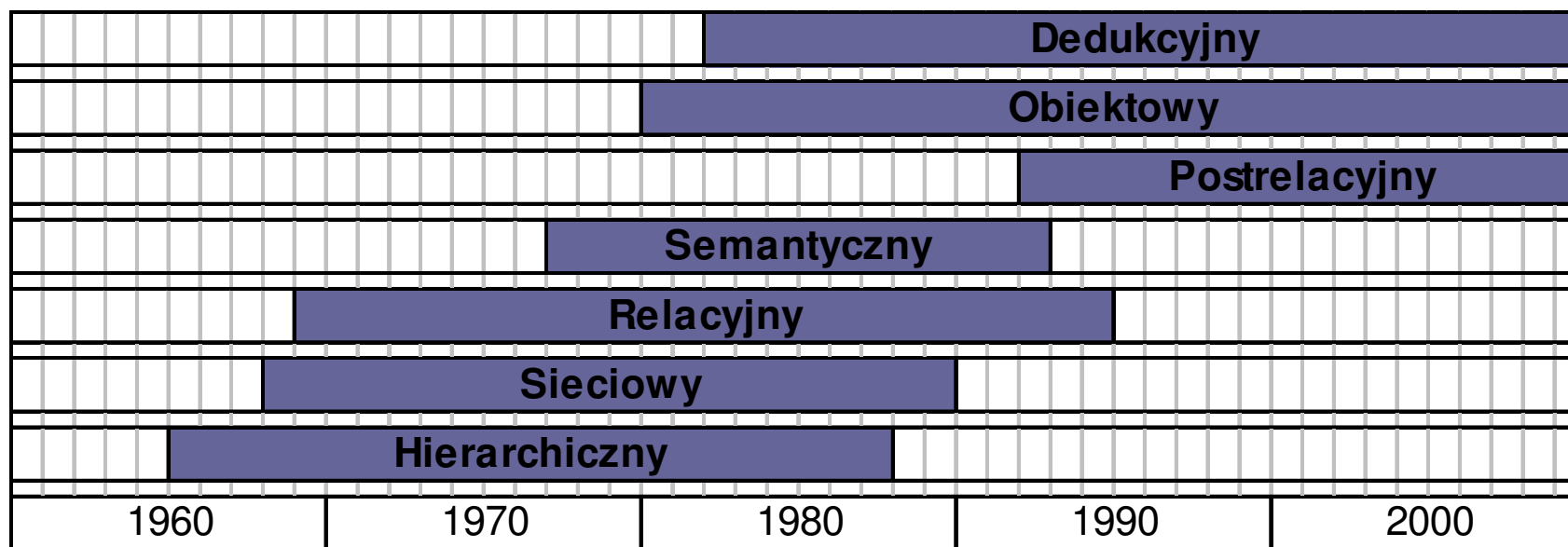
- Dane każdej organizacji podlegają nieustannym zmianom.
- W miarę stabilne pozostają jedynie ich:
 - » rodzaje
 - » sposób przechowywania i przetwarzania
- Modelowanie danych jest techniką organizowania i dokumentowania danych.
- Poprzez uogólnienie ich typów, cech i zależności między nimi można tworzyć modele danych. Modele danych można opracowywać na różnych poziomach abstrakcji czy szczegółowości. Najczęściej wyróżnia się:
 - » podstawowe modele danych, (konceptualne bądź logiczne), są ukierunkowane na potrzeby użytkownika, opisują dziedzinę przedmiotową, niezależnie od technicznego sposobu jego wdrożenia.
 - » wdrożeniowe modele danych dotyczą wdrożenia modelu danych w konkretnej technologii baz danych.

Cele modelowania danych

- Cele modelowania danych:
 - » Otrzymanie dokładnego **modelu potrzeb** informacyjnych przedsiębiorstwa,
 - » **Dekompozycja** i strukturalizacja problemu,
 - » Sformalizowanie opisu z wykorzystaniem języka graficznego – **jednoznaczność** i czytelność,
 - » Mechanizm efektywnej **komunikacji** pomiędzy analitykiem i użytkownikiem, pomiędzy analitykami systemu, a nawet pomiędzy użytkownikami,
 - » Poprawa jakości i efektywności projektowania bazy danych,
 - » Opis danych niezależny od struktur logicznych i fizycznych,
 - » **Niezależność** od implementacji pozwala na zastosowanie modelu do integracji istniejących baz danych,
 - » Podstawa do **zrozumienia** procesów realizowanych w przedsiębiorstwie i jego reorganizacji,
 - » Możliwość prezentacji **potrzeb informacyjnych** na różnym poziomie.

- **Model konceptualny** – spojrzenie na dane jako całość, model najbardziej stabilny, powinien on być podstawą, na której opierać się będzie przetwarzanie danych
- **Model wewnętrzny**, niskiego poziomu – opisuje sposób przechowywania danych w pamięci komputerów i przedstawia formaty rekordów czy ścieżki dostępu, modelami takimi są metody adresowania, struktury łańcuchowe i pierścieniowe

Historia rozwoju BD



- Modele użytkowe – stanowią podstawę do budowy systemu informatycznego:
 - » hierarchiczny
 - » sieciowy
 - » relacyjny
 - » obiektowy
 - » postrelacyjny

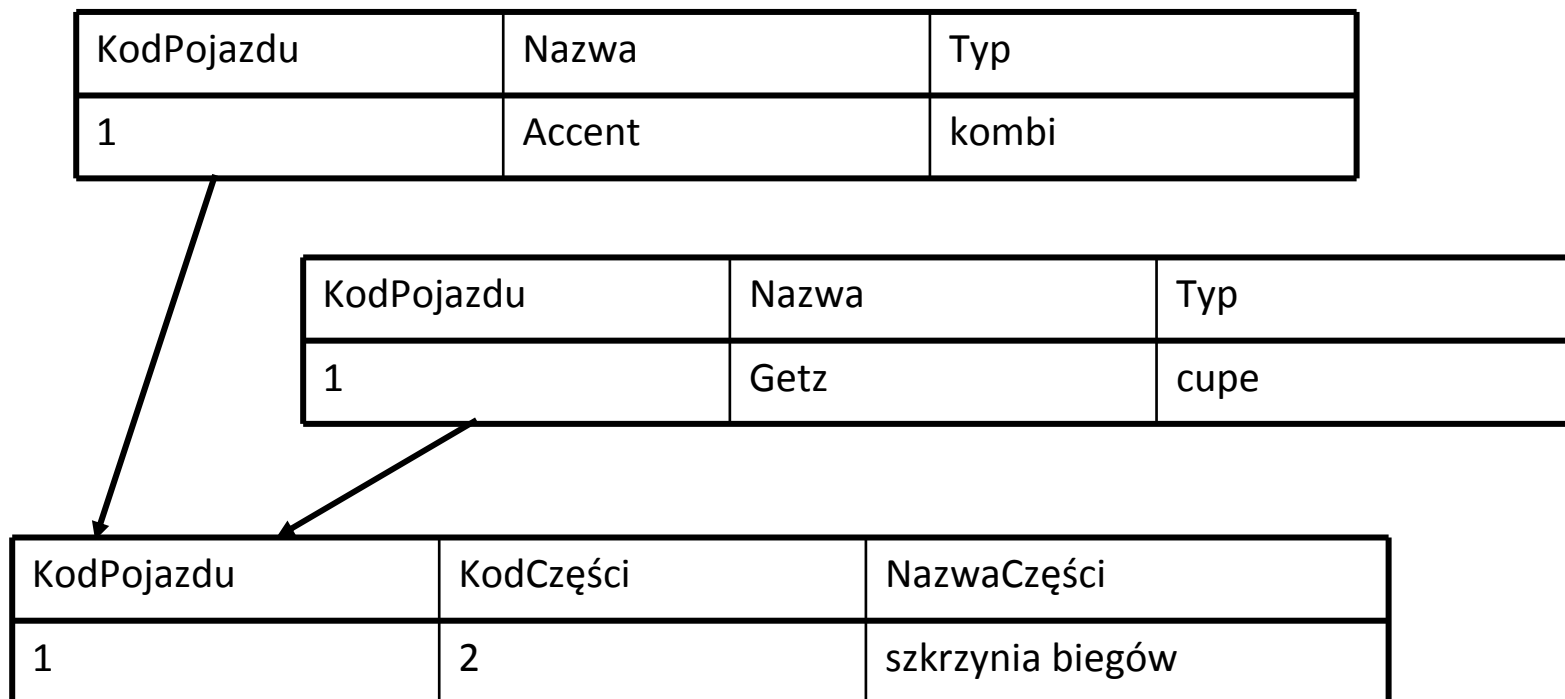
- Model obejmuje dwie struktury danych
 - » typy rekordów
 - » związki nadrzędny – podrzędny
- Każdy element zwany rekordem może uczestniczyć w roli podrzędnej w co najwyżej jednym powiązaniu rekordów, w roli nadrzędnej w dowolnej liczbie powiązań
- Rekord podrzędny nie może istnieć bez rekordu nadrzędnego
- Podmiotem operacji jest jeden rekord

Więzi w modelu hierarchicznym



- Model obejmuje dwie struktury danych
 - » typy rekordów
 - » typy kolekcji
- Każdy rekord może jednocześnie uczestniczyć w wielu powiązaniach rekordów
- Rekord taki może równocześnie i wielokrotnie wystąpić w roli nadrzędnej oraz w roli podrzędnej, powiązania realizowane są przez rekordy specjalne zwane łącznikami
- Podmiotem operacji jest jeden rekord

Więzi w modelu sieciowym



- Brak sprecyzowanej definicji obiektowych baz danych
- Model opiera się na takich pojęciach jak:
 - » klasa
 - » obiekt
 - » uogólnienie
 - » abstrakcja
 - » dziedziczenie
- Obiekty dysponują metodami

- Zbliżony do modelu obiektowego
- Skupia się na abstrakcji struktury, a nie na abstrakcji działania

- Oparty na logice formalnej
- Wykorzystywane elementy to
 - » predykaty
 - » argumenty
- Predykaty oraz argumenty tworzą asercję (zdanie), które może przyjmować wartość „prawda” lub „fałsz”
- Często oparty jest o język Datalog

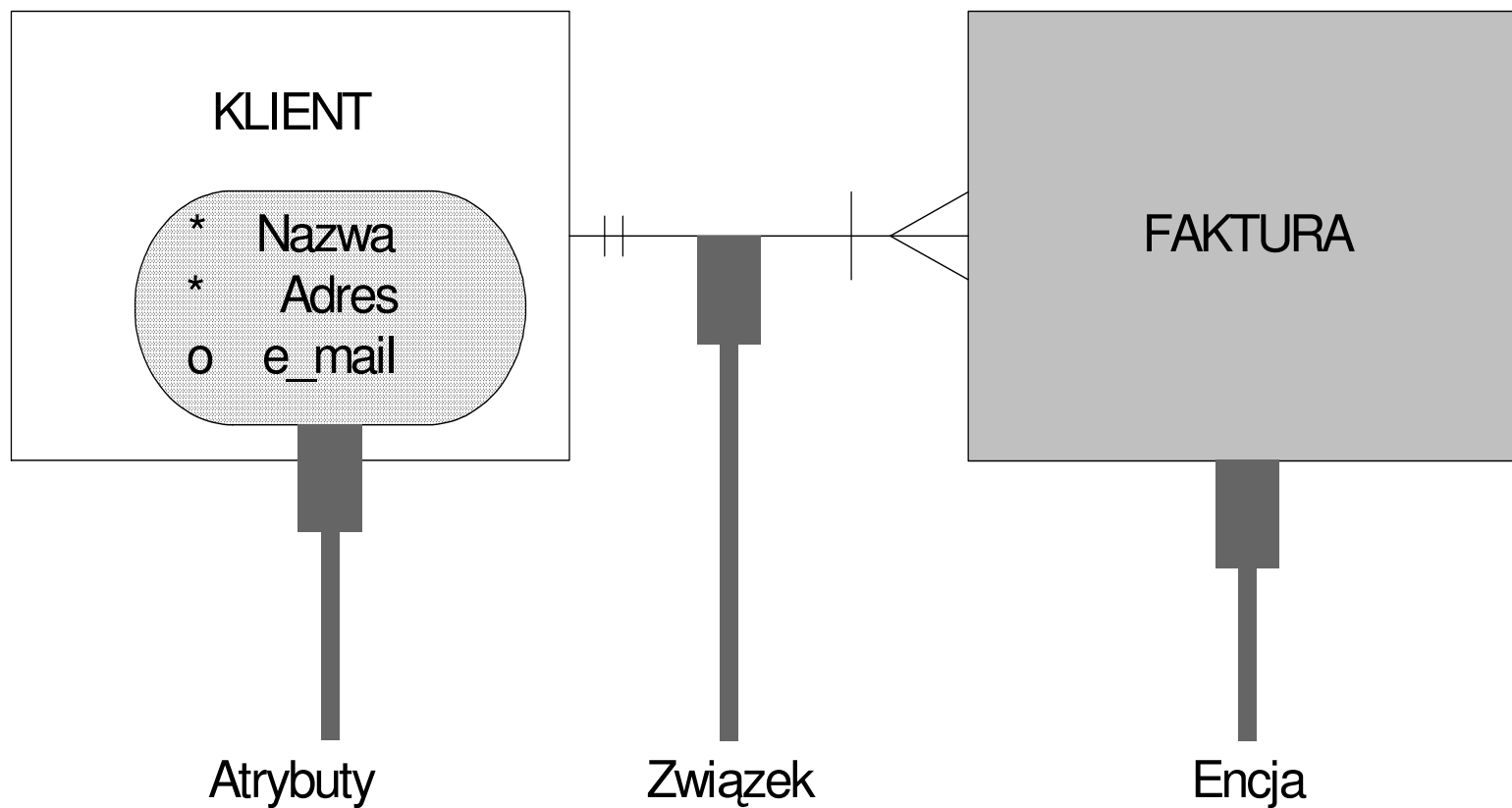
- Model relacyjny rozszerzony o elementy obiektowości
- Brak ścisłej definicji – za bazy realizujące model postrelacyjny przyjmuje się implementacje, które „już nie są relacyjne, ale jeszcze nie są obiektowe”

- Diagramy związku encji (Entity Relationship Diagrams) to metoda graficznego modelowania struktur danych oraz relacji między nimi
- Przedstawiają strukturę danych opisywanego systemu wraz z wszystkimi niezbędnymi atrybutami dla jego funkcjonowania.
- Modele danych można opracowywać na różnych poziomach szczegółowości.
- Modelowanie „z dołu do góry” (normalizacja) – konieczność zidentyfikowania całości zbioru danych przed projektowaniem
- Modelowanie „z góry do dołu” (modelowanie danych) – zbiór danych powstaje w trakcie projektowania
- Modelowanie semantyczne

Komponenty diagramu związków encji

| Komponent | Opis |
|------------------|---|
| Encja | <ul style="list-style-type: none">• Rzecz mająca znaczenie, rzeczywista lub wymyślona, o której informacje należy znać lub przechowywać. |
| Atrybut | <ul style="list-style-type: none">• Element informacji służący do klasyfikowania, identyfikowania, kwalifikowania, określania ilości lub wyrażania stanu encji. |
| Związek | <ul style="list-style-type: none">• Znaczący sposób, w jaki mogą być ze sobą powiązane dwie rzeczy tego samego typu lub różnych typów. |

Przykład prostego diagramu związków encji



- Encja (ang. entity) – jest to jednoznacznie identyfikowany składnik badanej rzeczywistości, o którym informacja jest lub może być zbierana i przechowywana.

- Przykładami encji są:
 - » PRACOWNIK,
 - » KLIENT,
 - » DOSTAWCA,
 - » ZAMÓWIENIE,
 - » MAGAZYN,
 - » KONTO itp.

- Uwaga: encje zazwyczaj opisuje się za pomocą rzeczowników lub wyrażeń rzeczownikowych w liczbie pojedynczej

Atrybut

- Atrybut - jest cechą, elementem charakteryzującym encje i związki w badanej dziedzinie przedmiotowej.

- Zestaw atrybutów, który jednoznacznie opisuje encję, nazywa się **wiązką atrybutów**.
- Wiązka powinna składać się, z co najmniej dwóch atrybutów opisujących daną encję. Szczególną rolę w zakresie atrybutów encji pełni **klucz**, zwany identyfikatorem. Pozwala on na jednoznaczne określenie wystąpienia encji.
- Jeśli używa się jednego atrybutu dla określenia encji, to mamy do czynienia z kluczem pojedynczym, a jeśli w tym celu używa się więcej niż jednego atrybutu, to z **kluczem złożonym**.

– Atrybut ma jedno z pięciu zadań:

- » identyfikować,
- » opisywać,
- » klasyfikować,
- » określać ilość,
- » wyrażać stan encji.



Rodzaje atrybutów

| Przykład | Przeznaczenie |
|--------------------------------|----------------------|
| numer zamówienia | identyfikacja |
| opis towaru | opis werbalny |
| typ towaru | klasyfikacja |
| ilość towaru w magazynie | określenie ilości |
| status płatności za zamówienie | wyrażenie stanu |

Wymagania dla atrybutu

- nazwa atrybutu musi być **unikalna** w ramach encji;
- atrybut musi być **obowiązkowy** (tzn., że wartość atrybutu musi być zawsze określona) lub **opcjonalny** (tzn., że atrybut nie musi mieć wartości). Symbolu „*” używa się dla atrybutu obowiązkowego, zaś symbolu „○” dla opcjonalnego;
- atrybut musi mieć format lub typ;

Przykładowe atrybuty

| Encja – STUDENT | Wystąpienia encji: | | | | | | | |
|---|--------------------|--------------------|--------|-----------------|------------------|---------------------|--|--|
| <table border="1" data-bbox="667 469 918 663"> <tr><td>STUDENT</td></tr> <tr><td># nr albumu</td></tr> <tr><td>* imię</td></tr> <tr><td>* nazwisko</td></tr> <tr><td>* data urodzenia</td></tr> <tr><td>* miejsce urodzenia</td></tr> </table>  | STUDENT | # nr albumu | * imię | * nazwisko | * data urodzenia | * miejsce urodzenia | <div data-bbox="1317 450 1536 619"> <p><i>Mat/123/04 Jan Kowalski 14-05-1990 Dobre Miasto</i></p> </div> <div data-bbox="1317 647 1536 817"> <p><i>Mat/345/04 Anna Nowak 21-05--1986 Dobre Miasto</i></p> </div> | |
| STUDENT | | | | | | | | |
| # nr albumu | | | | | | | | |
| * imię | | | | | | | | |
| * nazwisko | | | | | | | | |
| * data urodzenia | | | | | | | | |
| * miejsce urodzenia | | | | | | | | |
| Encja – SAMOCHÓD | Wystąpienia encji: | | | | | | | |
| <table border="1" data-bbox="698 1091 958 1318"> <tr><td>SAMOCHÓD</td></tr> <tr><td># nr rejestracyjny</td></tr> <tr><td>* typ</td></tr> <tr><td>* rok produkcji</td></tr> <tr><td>* cena</td></tr> <tr><td>* kolor</td></tr> <tr><td>* pojemność silnika</td></tr> </table>  | SAMOCHÓD | # nr rejestracyjny | * typ | * rok produkcji | * cena | * kolor | * pojemność silnika | <div data-bbox="1317 1040 1585 1225"> <p><i>OLX 2361 Nissan Almera 2000 55000 Czerwony 1,6 m³</i></p> </div> <div data-bbox="1317 1238 1572 1433"> <p><i>OM- 2388 Renault 2004 62000 Złoty 1,4 m³</i></p> </div> |
| SAMOCHÓD | | | | | | | | |
| # nr rejestracyjny | | | | | | | | |
| * typ | | | | | | | | |
| * rok produkcji | | | | | | | | |
| * cena | | | | | | | | |
| * kolor | | | | | | | | |
| * pojemność silnika | | | | | | | | |

- Związek stanowi naturalne powiązanie pomiędzy dwoma lub więcej encjami w badanej dziedzinie przedmiotowej.
- W identyfikowaniu i modelowaniu związków encji bierze się pod uwagę następujące cechy:
 - » stopień związku (liczebność związku)
 - » opcjonalność (uczestnictwo encji).

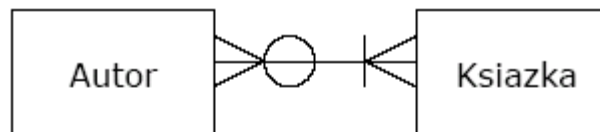
- oznacza stosunek ilościowy między liczebnością **wystąpień** poszczególnych encji, uczestniczących w danym związku,
- mówi o tym, ile wystąpień encji jednego rodzaju jest powiązanych z iloma wystąpieniami encji innego rodzaju

Przykłady związków encji

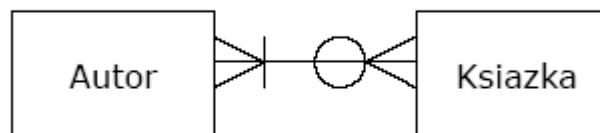
| Stopień związku | Przykład | Znaczenie |
|----------------------|---------------------|--|
| 1:1 | Dziekan- Wydział | <ul style="list-style-type: none"> Każde wystąpienie encji Dziekan jest powiązane tylko z jednym wystąpieniem encji Wydział. Zatem jeden Dziekan kieruje jednym Wydziałem |
| 1:m 1: wiele | Wydział- Student | <ul style="list-style-type: none"> Każde wystąpienie encji Wydział powiązane jest jednym lub wieloma wystąpieniami encji Student, przy czym każde wystąpienie encji Student powiązane jest tylko jednym wystąpieniem encji Wydział. Zatem Wydział posiada wielu Studentów, natomiast Student studiuje wyłącznie na jednym Wydziale |
| m:n Wiele : wiele | Książka - Autor | <ul style="list-style-type: none"> Każde wystąpienie encji Książka powiązane jest z wieloma wystąpieniami encji Autor i odwrotnie każde wystąpienie encji Autor powiązane jest z wieloma wystąpieniami encji Książka. Jest to sytuacja, gdzie Książka może być napisana przez jednego lub wielu autorów i jeden Autor jest podpisany pod jednym lub wieloma tytułami Książek. |

Formy zapisu związku

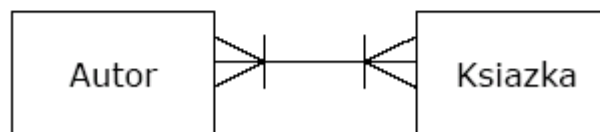
autor musi mieć napisaną
co najmniej 1 książkę



książka musi być napisana
przez co najmniej 1 autora



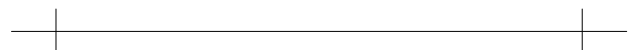
autor musi mieć napisaną
co najmniej 1 książkę
książka musi być napisana
przez co najmniej 1 autora



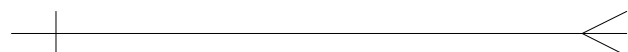
- dotyczy zaangażowania encji w związek,
- z uwagi na tę cechę wyróżnia się dwa typy związków:
 - » **wymagane (obowiązkowe)** – zachodzi wówczas, jeśli wszystkie wystąpienia encji muszą uczestniczyć w związku;
 - » **opcjonalne** - zachodzi wówczas, jeśli istnieje, co najmniej jedno wystąpienie encji, które nie uczestniczy w związku.

Cechy związków encji (notacja Martina)

Stopień związku



jeden - do - jednego

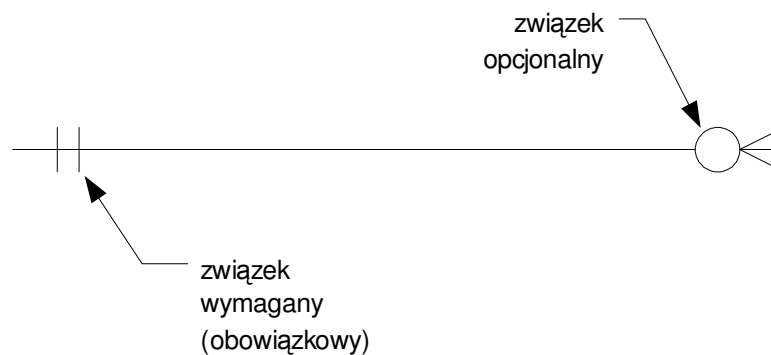


jeden - do - wielu



wiele - do - wielu

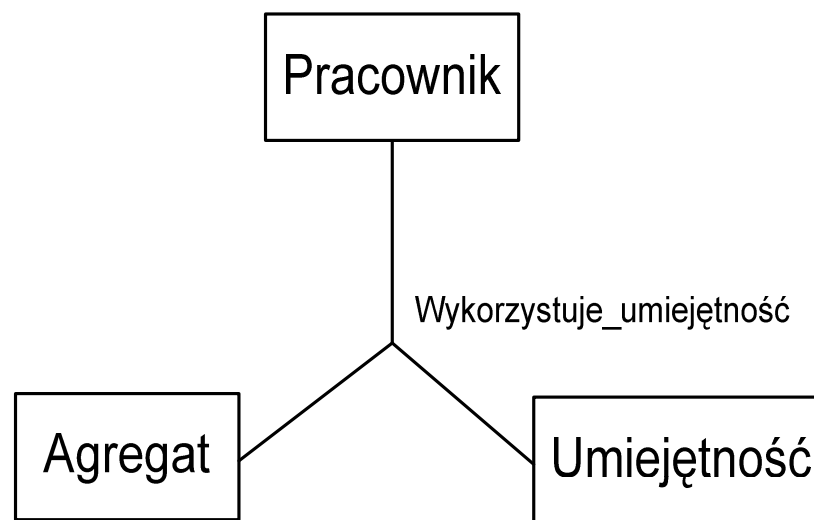
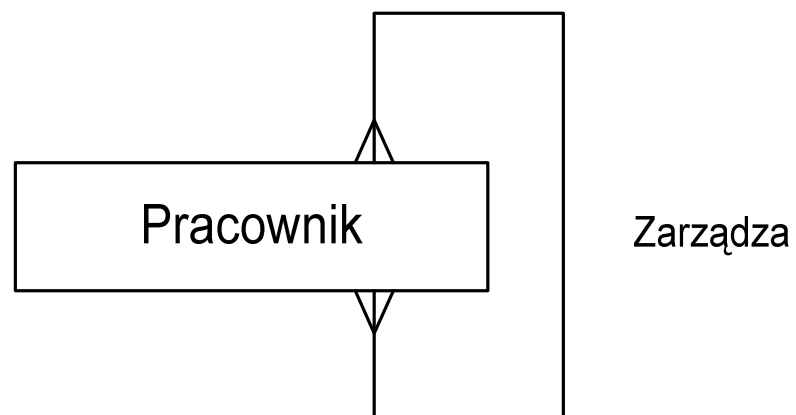
Typ związku (opcjonalność)



- Liczebność (stopień)
 - » 1:1
 - » 1:M
 - » M:N
 - » Liczebność min/max
- Uczestnictwo (opcjonalność)
- Zawieranie i wykluczanie

Diagramy Związków Encji

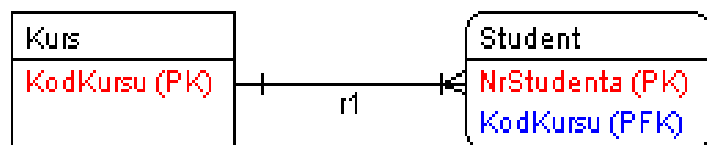
- Związki rekurencyjne (jednoargumentowe)
- Związki trójargumentowe – rozbicie na dwie osobne relacje powoduje utratę informacji
- Role
- Atrybuty związków (możliwa konwersja do nowego zbioru encji)



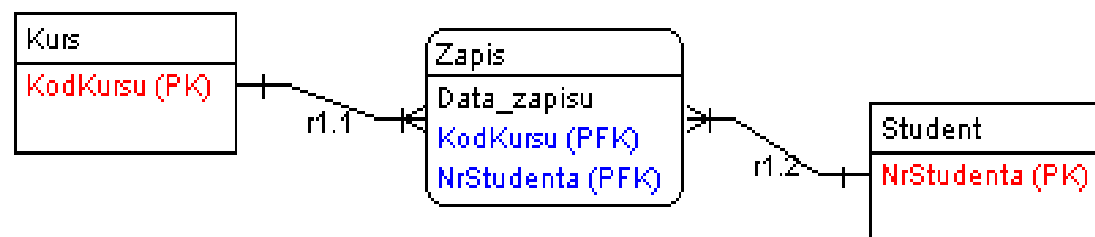
- Wyjątkowe zbiory encji zawierające dodatkowe/specjalizowane atrybuty/związki
- Zbiory encji łączone są z podklasami związkami „isa”
- Kreskówka **is a** Film
- Kryminał **is a** Film

ERD – Modelowanie upływu czasu

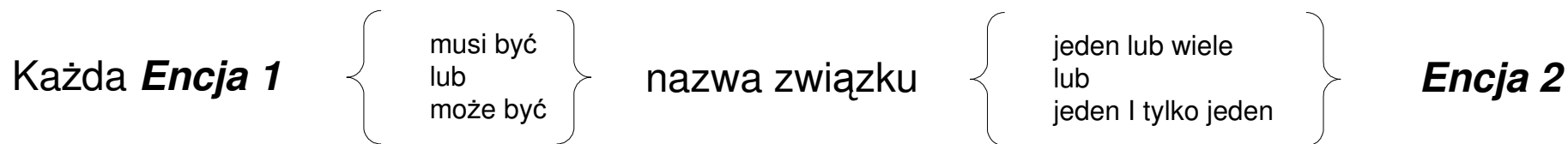
[1.1]



[1.1]



Reguły czytania związków encji



Przykład:



Każdy **Klient** może złożyć jedno lub wiele **Zamówień**.

Każde **Zamówienie** musi być zlecone przez jednego i tylko jednego **Klienta**.

- tworzenie diagramu związków encji najlepiej rozpocząć od wskazania **encji** oraz określić **związki** między encjami występującymi w danej dziedzinie przedmiotowej.
- istnieje kilka najbardziej rozpowszechnionych notacji graficznych diagramu ERD, należą do nich notacje:
 - Chena,
 - Bachmana,
 - Martina,
 - Shlaer-Mellora.
- ponieważ w zasadzie notacje te są równoważne i różnią się jedynie wyglądem symboli graficznych, do opisu wybrano najbardziej rozpowszechnioną **notację Martina**.

Typy encji ERD (notacja Martina)



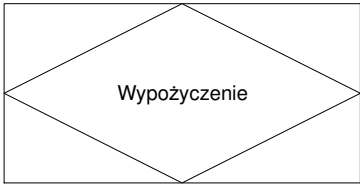
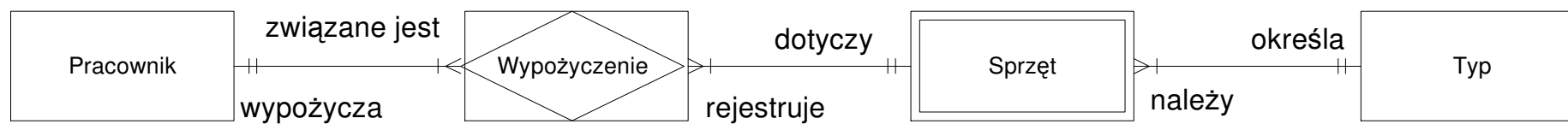
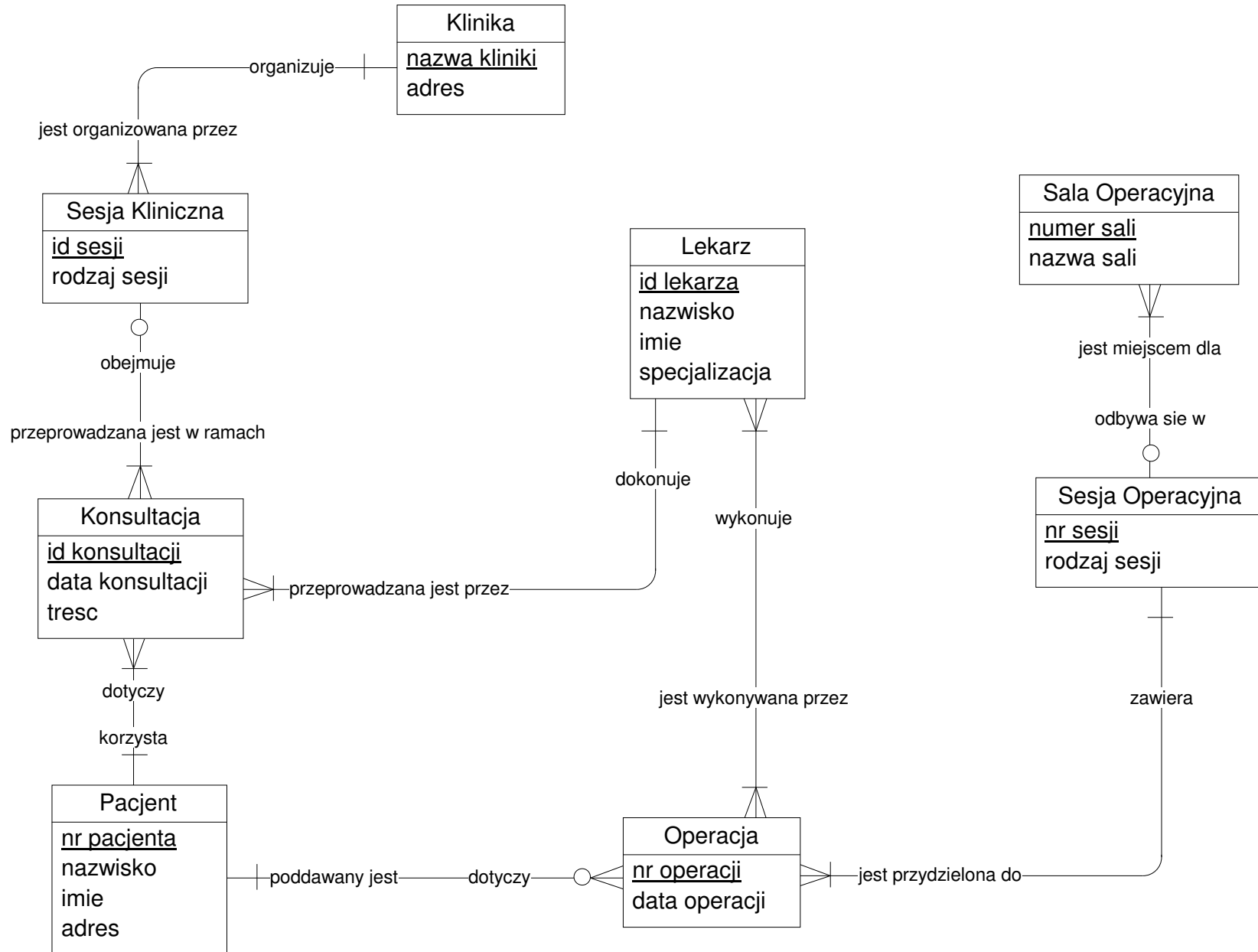
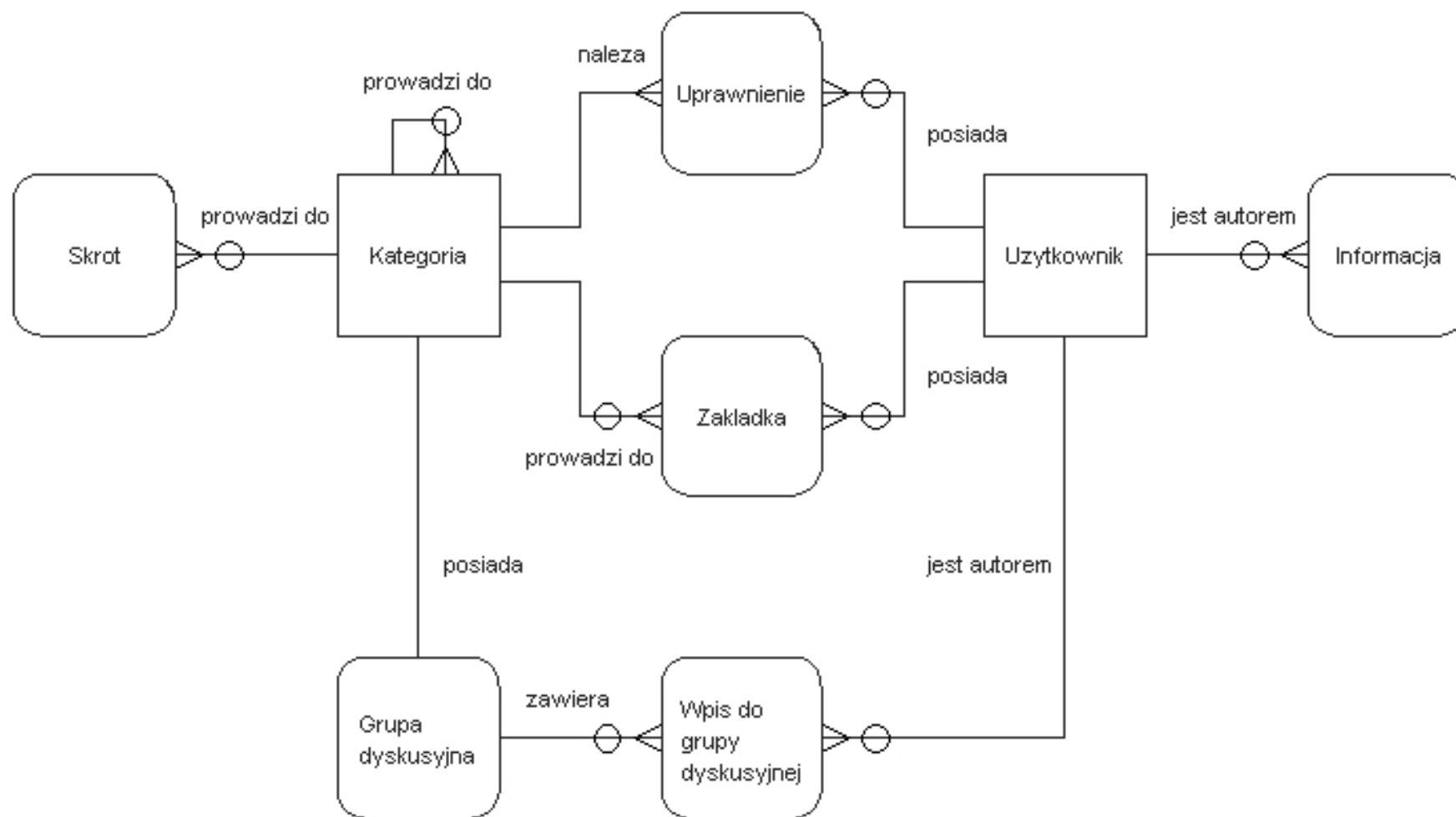
| Typy encji | Notacja Martina |
|--|---|
| <p>Encja regularna – oznacza dowolny znaczący element, o którym informacja powinna być znana albo utrzymywana (częściowe uczestnictwo w związku).</p> |  |
| <p>Encja słaba – jest to encja, która może istnieć tylko wtedy, gdy jest związana z innymi encjami lub też nie posiada własnych atrybutów kluczowych (całkowite uczestnictwo w związku)</p> |  |
| <p>Encja – obiekt asocjacyjny – przechowuje informacje o związku pomiędzy dwiema encjami.</p> |  |

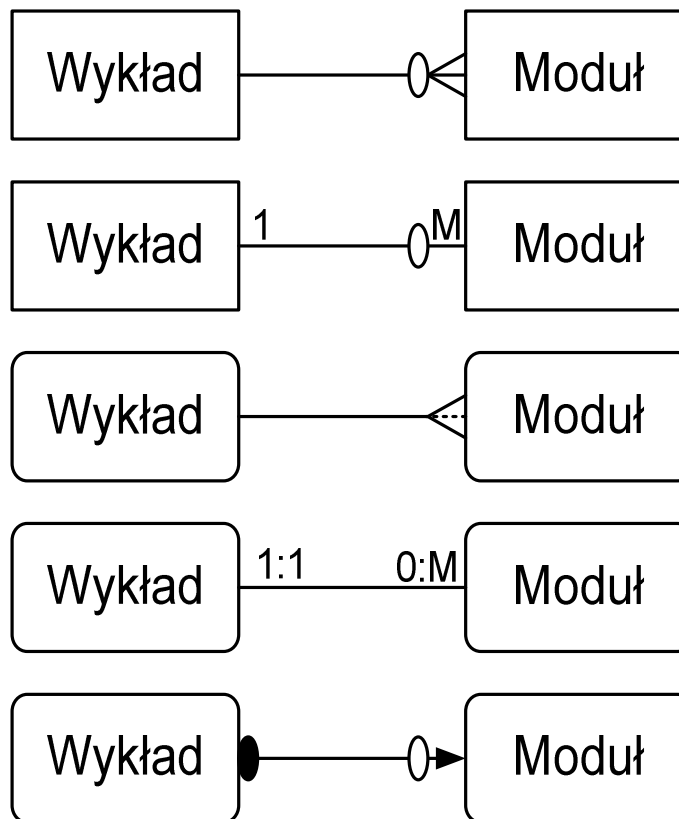
Diagram ERD z wyróżnionymi typami encji







Diagramy Związków Encji



- Konceptualny model danych, którego odzwierciedleniem są diagramy ERD, przekształcany jest w jeden z modeli baz danych:
 - » relacyjny,
 - » sieciowy,
 - » hierarchiczny.

Terminologia relacyjna

| Pojęcie | Opis |
|-----------------------------|---|
| Relacja | <ul style="list-style-type: none"> Jest to podzbiór iloczynu kartezyjskiego reprezentowany przez zbiór krotek. Reprezentacją relacji jest tablica. |
| Krotka | <ul style="list-style-type: none"> Oznacza wiersz tablicy. Reprezentacją krotki w tablicy jest rekord. |
| Atrybut | <ul style="list-style-type: none"> Oznacza kolumnę tablicy (a dokładniej są to różne wystąpienia tego samego atrybutu). Reprezentacją atrybutu w tablicy jest pole. |
| Stopień relacji | <ul style="list-style-type: none"> Liczba typów atrybutów w relacji. |
| Liczebność relacji | <ul style="list-style-type: none"> Liczba krotek w relacji. |
| Klucz główny | <ul style="list-style-type: none"> Kolumna lub kombinacja kolumn, których wartości jednoznacznie identyfikują wiersze w tablicy. |
| Klucz obcy | <ul style="list-style-type: none"> Kolumna lub kombinacja kolumn, których wartości określają klucz główny innej tablicy. |
| Dziedzina (atrybutu) | <ul style="list-style-type: none"> Lista dostępnych wartości atrybutu, wszystkie tego samego typu. |

- każda encja staje się tablicą, której nazwa jest zazwyczaj nazwą encji w liczbie mnogiej;
- każdy atrybut staje się komuną, a jego nazwa odpowiednio nazwą tej kolumny. Natomiast właściwości atrybutu stają się odpowiadającymi im właściwościami w projekcie danych. Atrybuty obowiązkowe stają się kolumnami NOT NULL (co oznacza, że nie jest możliwe by wartość kolumny przyjmowała wartość NULL);
- unikalny identyfikator encji staje się kluczem głównym tabeli;
- każdy związek jest przekształcany w dwa obiekty. Kolumnę klucz obcego, zgodną z kluczem głównym (lub unikalnym) tabeli, której dotyczy. Dziedziczy ona typ i rozmiar danego klucza głównego.

Przekształcanie encji

Logiczny model danych

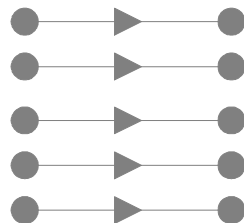
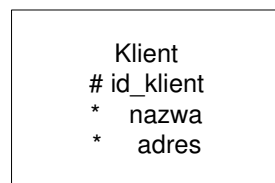
Encja

Klient

Atrybuty

nazwa

unikalny identyfikator



Relacyjny model danych

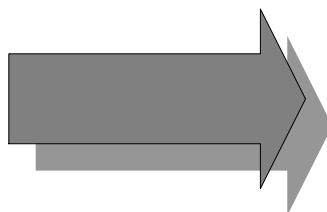
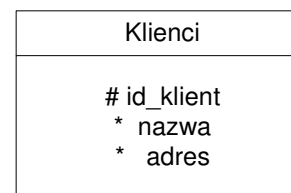
Tabela

Klienci

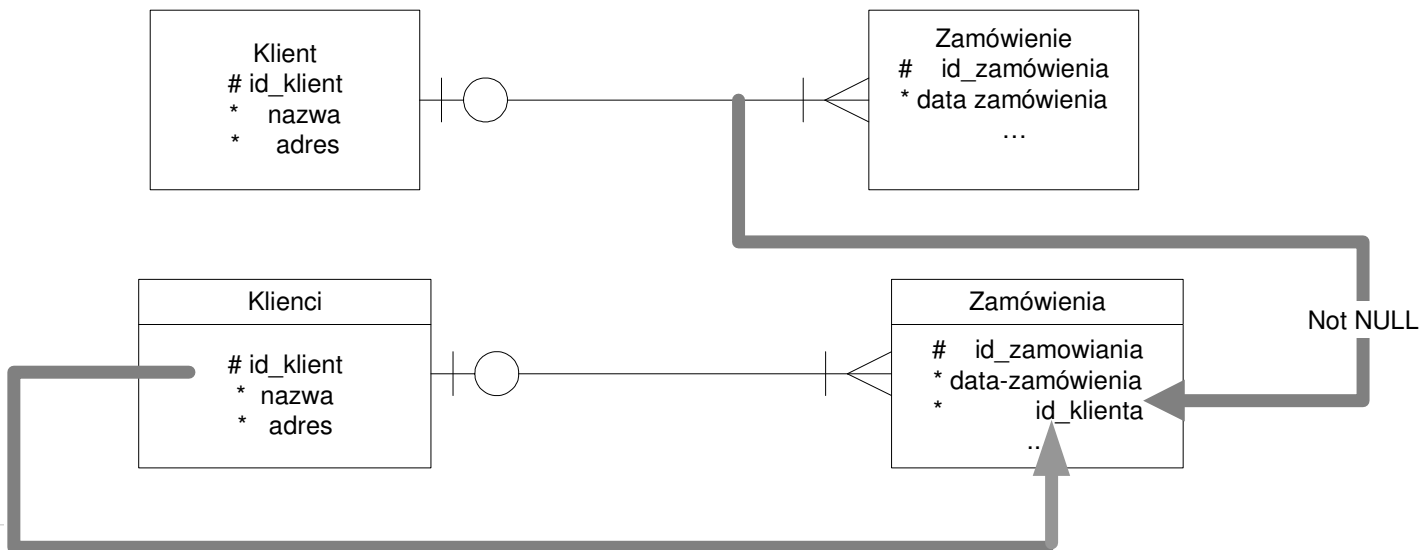
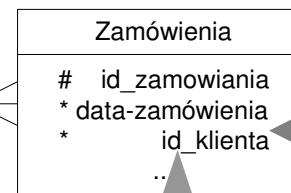
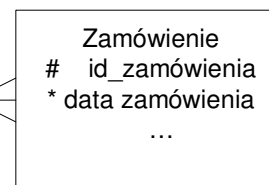
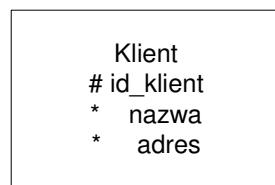
Kolumny

nazwa

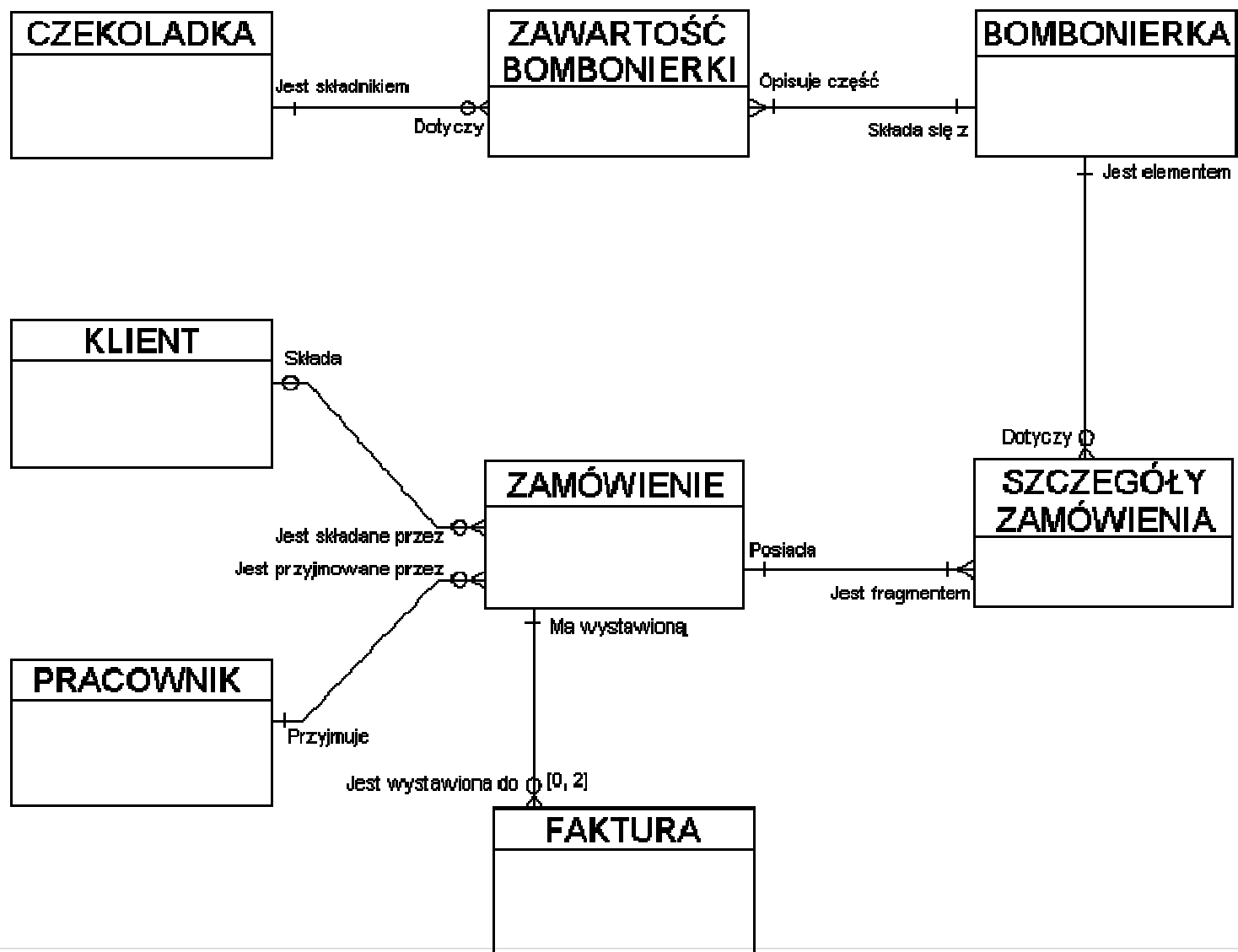
klucz główny



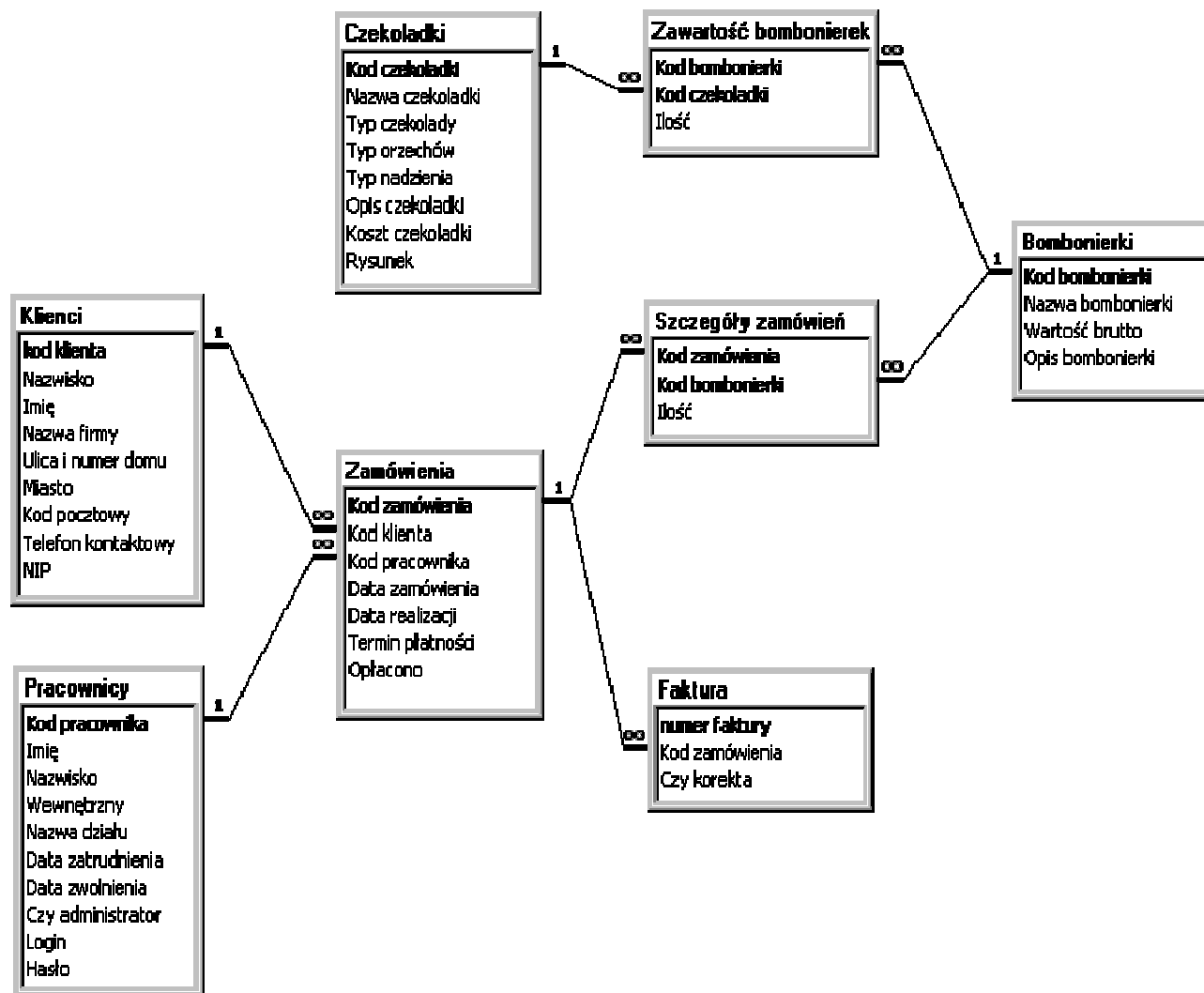
Przekształcanie związków



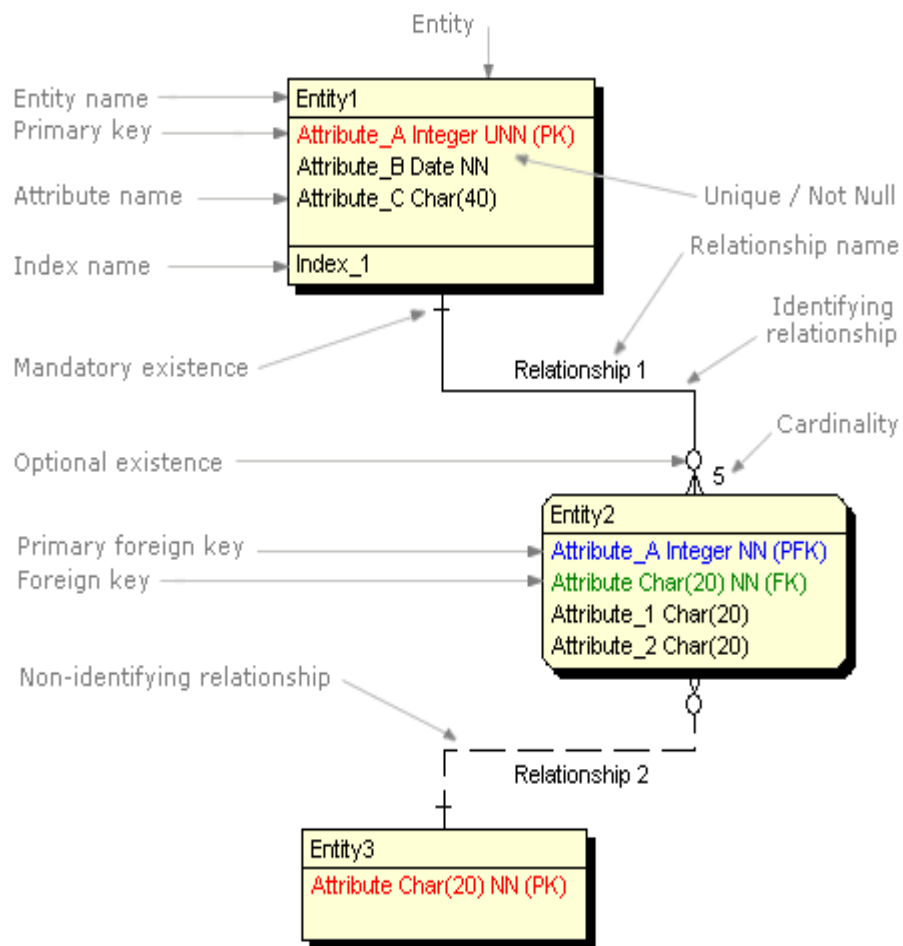
Przekształcenie diagramu ERD



Przekształcenie diagramu ERD



Toad Data Modeler - ERD



Toad Data Modeler - ERD

Parent: mandatory Child: mandatory



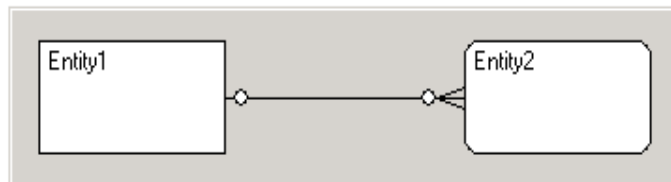
Parent: mandatory Child: optional



Parent: optional Child: mandatory



Parent: optional Child: optional



Identifying relationship



Non-identifying relationship



Informative relationship

