

Projektowanie. Normalizacja.

Krzysztof Regulski

WIMiP, KISiM,

regulski@agh.edu.pl

B5, pok. 408

Etapy projektowania baz danych:

1. **Specyfikacja wymagań użytkownika** - określenie zjawisk, dostępności i użyteczności danych, ich formatu i sposobów obliczeń, cele, zakres i kontekst systemu
2. **Projektowanie koncektualne** - projektowanie schematu E–R bazy. Użycie modelu E–R wpływa również na realizację pozostałych faz.
3. **Specyfikacja wymagań funkcjonalnych** - dokładny opis wymagań klienta i wszystkich przyszłych użytkowników systemu
4. **Projektowanie logiczne i fizyczne**
5. **Implementacja**

Projektowanie bazy danych:

– Projektowanie logicznej struktury bazy:

- » **Etap I:** określenie encji i zdefiniowanie atrybutów opisujących encje
 - przyporządkowanie encji do zjawisk
 - standaryzacja nazw i formatów
 - identyfikacja źródeł danych
- » **Etap II:** określenie związków między encjami
 - identyfikacja typu związków (relacji) (1-1, 1-M, N-M)
- » **Etap III:** normalizacja relacji
 - obniżenie redundancji i wyeliminowanie anomalii (usuwania, wstawiania i aktualizacji)

– Projektowanie fizycznej struktury bazy:

- » nałożenie struktury logicznej na fizyczne urządzenia

Modelowanie schematu relacji

- Projekt bazy danych polega na znalezieniu właściwych schematów relacji tworzących bazę danych
- **Niewłaściwy projekt** (niewłaściwe schematy) mogą prowadzić do
 - » Redundancji (powtarzania informacji).
 - » Niemożliwości reprezentowania pewnych informacji.
 - » Anomalii związanych z operowaniem danymi (głównie modyfikacją danych)
- **Cele:**
 - » Unikanie redundancji danych
 - » Zapewnienie reprezentowania związków między danymi
 - » Zachowanie warunków integralności (umożliwienie kontroli warunków integralności podczas modyfikacji danych).

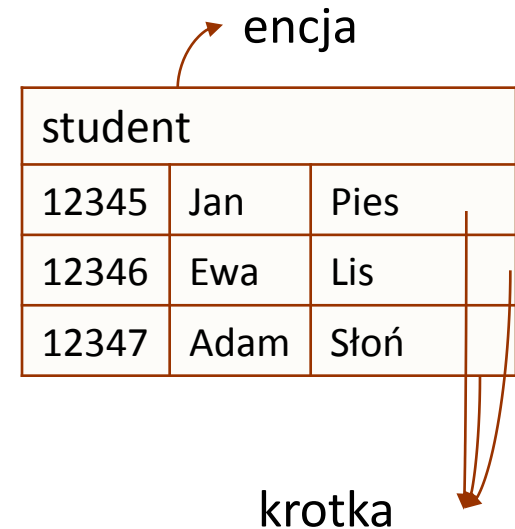
Encja (zbiory encji):

- **Encja (*entity*)** - model rzeczy, osób, zjawisk, pojęć itp., o których chcemy przechowywać informacje, które mają tożsamość (są rozróżnialne)

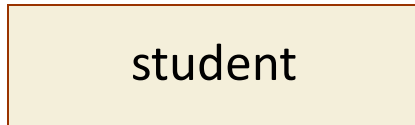
Np. konkretna osoba, firma, zdarzenie

- **Zbiór encji (*entity set*)** – zbiór (klasa) obiektów, które są tego samego typu (mają wspólne własności), inaczej: zbiór wystąpień (instancji) tej klasy

Np. zbiór osób, zbiór firm, zbiór określonych zdarzeń



W diagramach E-R zbiory encji oznaczane są **prostokątami:**



student

Atrybuty:

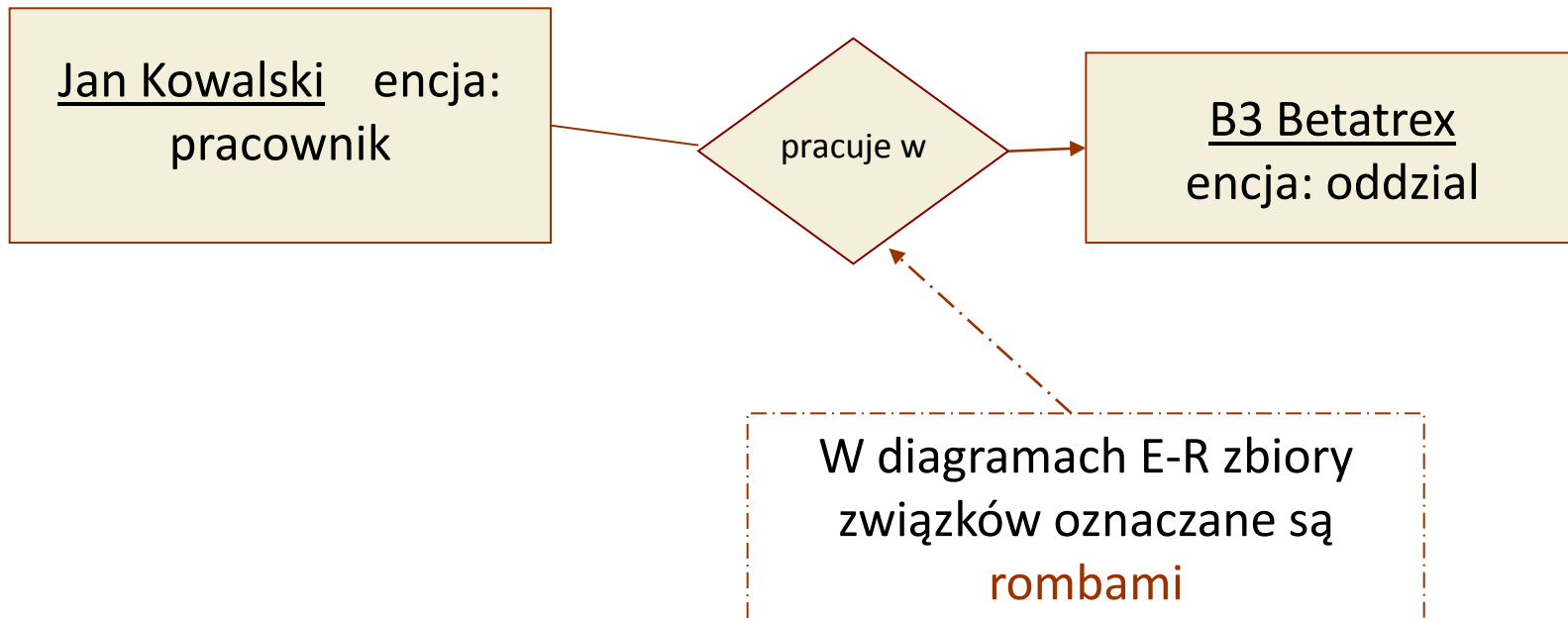
- Zbiór encji (i każda z nich) jest opisywany przez **zestaw atrybutów**, które odpowiadają własnościom posiadanym przez wszystkie obiekty reprezentowane przez ten zbiór. Np.:

student = (nr_albumu, imie, nazwisko)

- **Dziedzina atrybutu** – zbiór dopuszczalnych wartości, jakie może przyjmować dany atrybut na danym zbiorze encji. Dziedzina uzupełniona jest o wartość **NULL** oznaczającą, że encja nie posiada wartości atrybutu lub jest on nieznan.
- **Atrybuty:**
 - » proste vs. pochodne (mogą być obliczone na podstawie innych informacji w bazie),
 - » proste vs. złożone.

Zbiory związków:

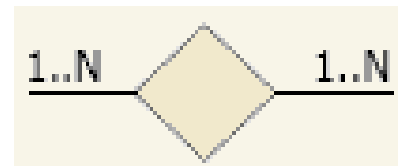
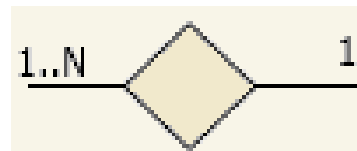
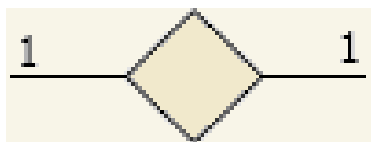
- **Związek** (*relationship*) reprezentuje relację pomiędzy pewnymi encjami (obiektami). Np.:



Stopień przyporządkowania (krotność związków):

- **Stopień przyporządkowania** zbiorów encji do związku (relacji) określony jest przez ilość encji z każdego zbioru, które są ze sobą w tym związku powiązane.
- **Rodzaje związków binarnych** ze względu na stopień przyporządkowania:
 - *jeden do jeden*
 - *jeden do wielu*
 - *wiele do wielu*
- **W diagramie E-R** rodzaje związków są oznaczane w następujący sposób:
 - (grot strzałki) oznacza stopień przyporządkowania *jeden*
 - (brak grotu) – *wielu*

lub:



Ograniczenia - warunki integralności

- W ramach projektowania baz danych specyfikuje się pewne ograniczenia ilościowe i zależności pomiędzy encjami lub związkami, które mają zachodzić **zawsze**.
- Ograniczenia te i zależności określają **warunki integralności** rozważanej bazy i są następnie sprawdzane i dotrzymywane.
- Do najistotniejszych warunków ogólnych, specyfikowalnych w modelu E-R należą:
 - » stopień przyporządkowania zbiorów encji do związku,
 - » zależność występowania encji z różnych zbiorów,
 - » całkowitość vs. częściowość udziału zbioru encji w zbiorze związków.

Klucz:

- **Kluczem relacji** nazywamy taki zbiór atrybutów tej relacji, których kombinacje wartości jednoznacznie identyfikują każdą krotkę tej relacji a żaden podzbiór tego zbioru nie posiada tej własności. W kluczu nie może zawierać się wartość NULL.
- Klucz jest **kluczem prostym**, jeżeli powyżej opisany zbiór jest jednoelementowy - w przeciwnym razie mówimy o **kluczu złożonym**.
- W ogólności, w relacji można wyróżnić wiele kluczy, które nazywamy **kluczami potencjalnymi**. Wybrany klucz spośród kluczy potencjalnych nazywamy **kluczem głównym**.

- Minimalny **zbiór identyfikujący**
- Taki zbiór atrybutów relacji, których kombinacje wartości **jednoznacznie identyfikują** każdą krotkę tej relacji a **żaden podzbiór** tego zbioru nie posiada tej własności
- W kluczu nie może zawierać się wartość Null

- Klucz jest **kluczem prostym**, jeżeli powyżej opisany zbiór jest jednoelementowy - w przeciwnym razie mówimy o **kluczu złożonym**
- W ogólności, w relacji można wyróżnić wiele kluczy, które nazywamy **kluczami potencjalnymi**. Wybrany klucz spośród kluczy potencjalnych nazywamy **kluczem głównym (Primary Key PK)**

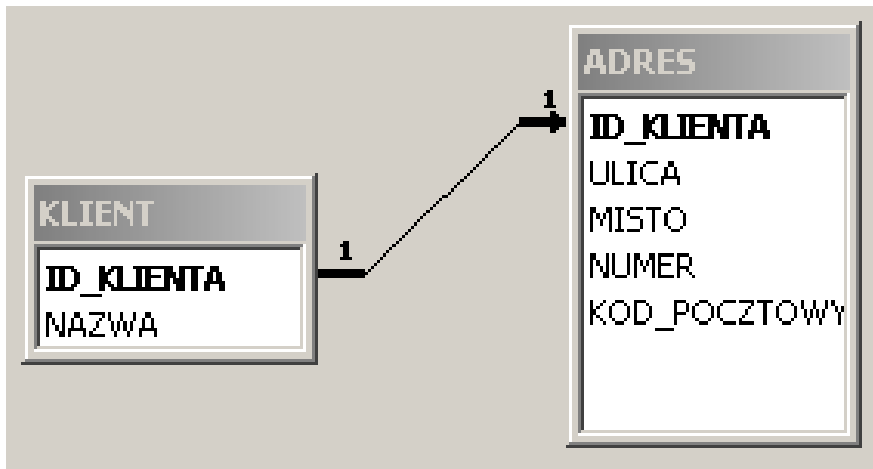
- Atrybut ***B*** relacji ***R*** jest funkcjonalnie zależny od atrybutu ***A*** jeżeli dowolnej wartości ***a*** atrybutu ***A*** odpowiada nie więcej niż jedna wartość ***b*** atrybutu ***B***

$$A \rightarrow B$$

- **Więź** (*relationship*) to powiązanie pomiędzy parą tabel (relacji).
- Istnieje ona wtedy, gdy dwie tabele są połączone przez **klucz podstawowy i klucz obcy**. Każda więź jest opisywana przez typ więzi istniejący między dwoma tabelami, typ uczestnictwa oraz stopień uczestnictwa tych tabel.

Typy więzi

- **jeden-do-jednego** (jeżeli pojedynczemu rekordowi z pierwszej tabeli przyporządkowany jest najwyżej jeden rekord z drugiej tabeli i na odwrót)



Edytowanie relacji [?] [X]

Tabela/Kwerenda: KLIENT Pokrewna tabela/kwerenda: ADRES

ID_KLIENTA	ID_KLIENTA

Wymuszaj więzy integralności
 Kaskadowo aktualizuj pola pokrewne
 Kaskadowo usuń rekordy pokrewne

Typ relacji: Jeden-do-jedne

OK
Anuluj
Typ sprzężenia..
Utwórz nowe..

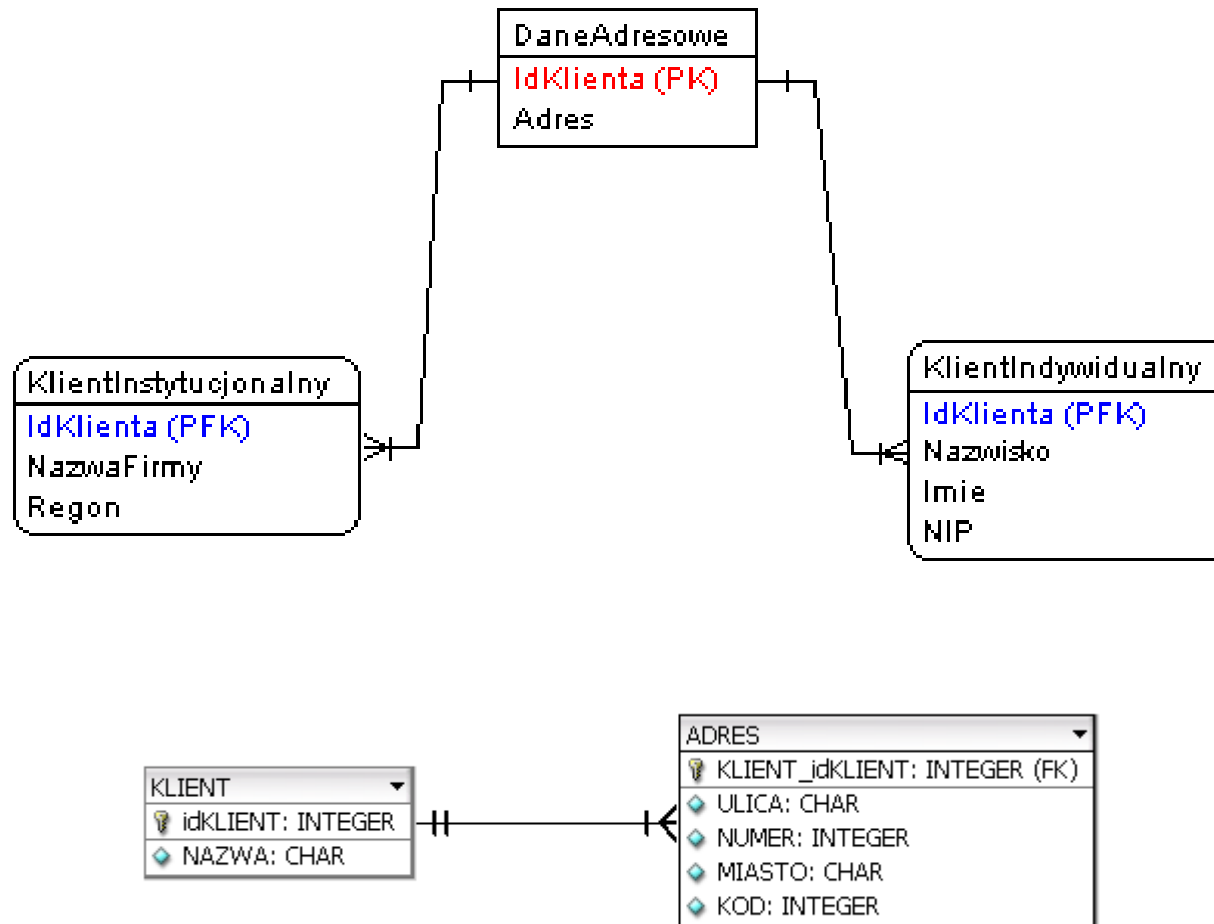
Właściwości sprzężenia [?] [X]

- 1: Zawiera tylko te wiersze, w których sprzężone pola z obu tabel są równe.
- 2: Uwzględnia **WSZYSTKIE** rekordy z 'KLIENT' i tylko te rekordy z 'ADRES', dla których sprzężone pola są równe.
- 3: Uwzględnia **WSZYSTKIE** rekordy z 'ADRES' i tylko te rekordy z 'KLIENT', dla których sprzężone pola są równe.

OK
Anuluj

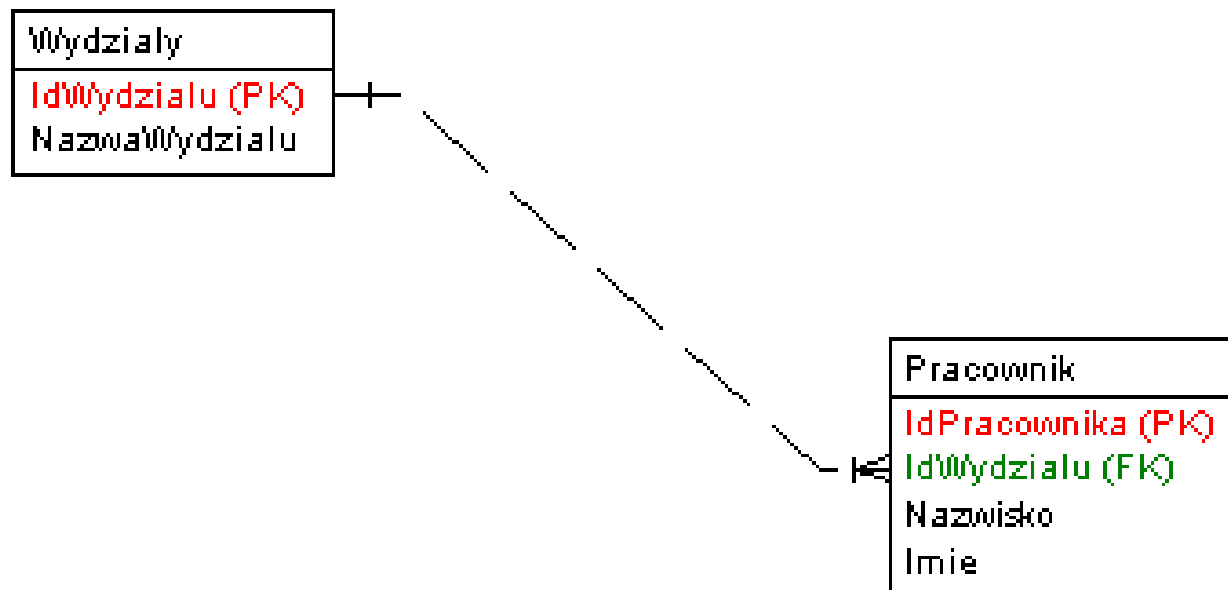
Więź jeden-do-jednego

[1.1]



Typy więzi

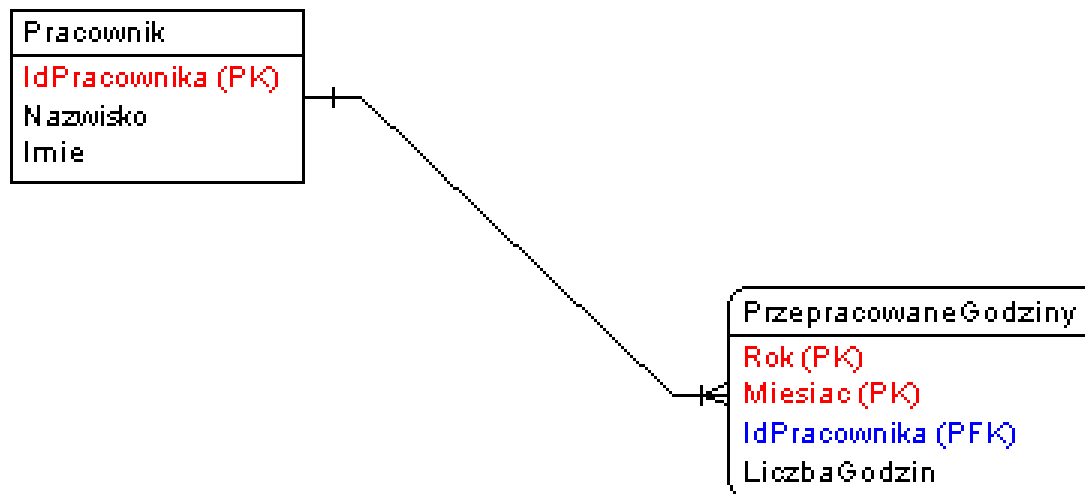
- **jeden-do-wielu** (jeżeli pojedynczemu rekordowi z pierwszej tabeli może odpowiadać jeden lub więcej rekordów z drugiej, ale pojedynczemu rekordowi z drugiej tabeli odpowiada najwyżej jeden rekord z tabeli pierwszej)



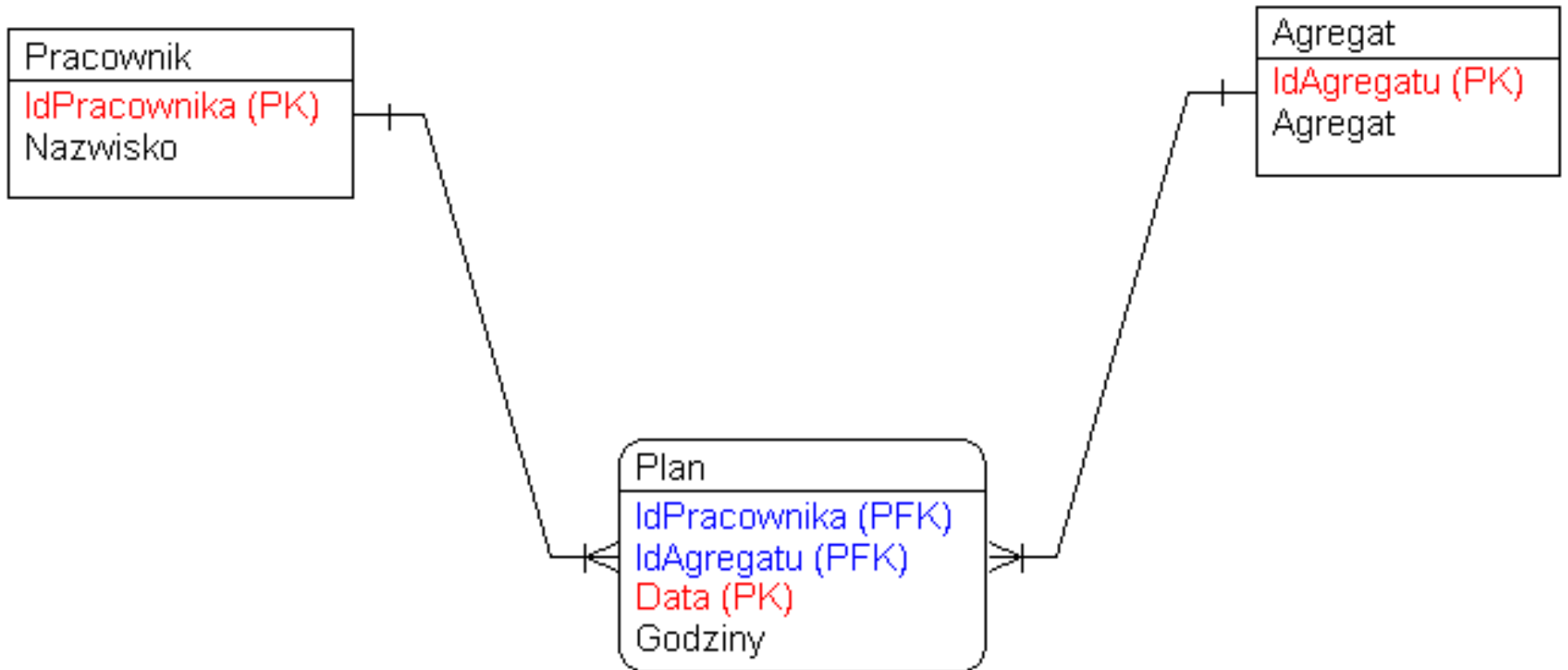
Więzi identyfikujące

- **Klucz obcy**, który jest składnikiem złożonego klucza głównego w relacji zależnej określany jest mianem **klucza obcego głównego (Primary Foreign Key)** a tak zbudowana więź jest **więzią identyfikującą**

[1.1]



Więź wiele-do-wielu

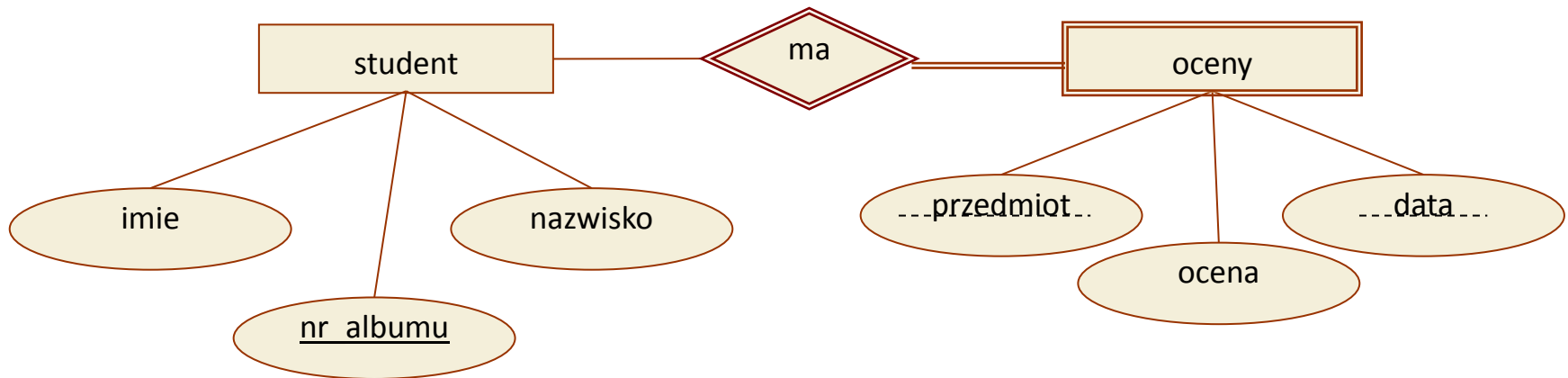


Zbiory słabe encji:

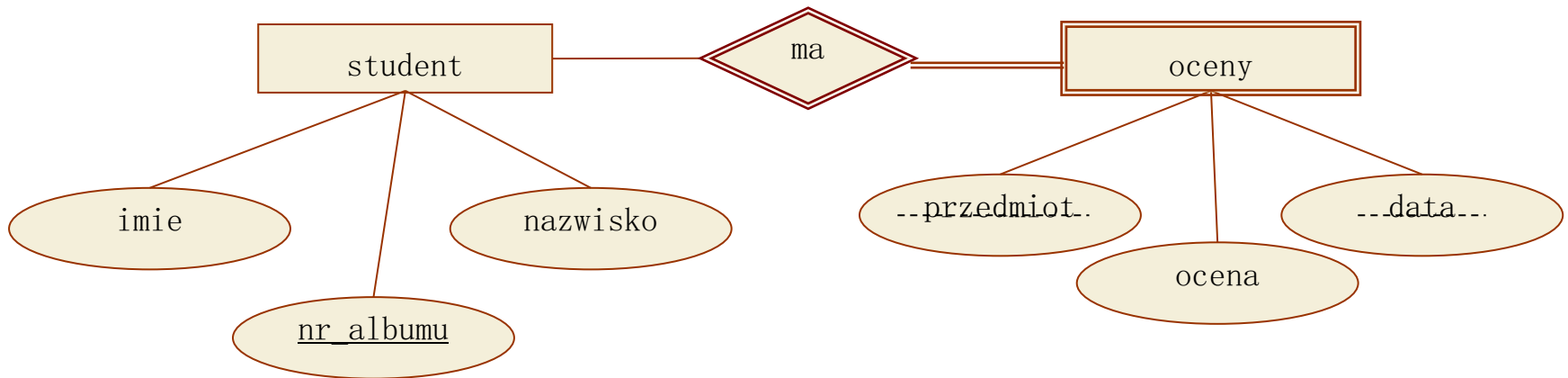
- Zbiór encji, który nie posiada klucza głównego (nie ma odpowiednich atrybutów), jest nazywany **zbiorem słabym encji** (*weak entity set*) w przeciwieństwie do zbioru silnego - (*strong entity set*).
- **Informacyjność** zbioru słabych encji jest uwarunkowana istnieniem innego silnego zbioru encji, który pozostaje ze słabym w związku *jeden do wielu*. Encje tego zbioru silnego dominują również nad encjami słabego w sensie zależności występowania.
- W zbiorze słabych encji występować musi wówczas **klucz częściowy** (*partial key*), który pozwala na rozróżnienie (częściowe) jego encji względem dominującego zbioru silnego.
- **Kluczem głównym** zbioru słabego encji jest kompozycja klucza głównego dominującego zbioru silnego encji i klucza częściowego zbioru słabego.

Zbiory słabe encji:

- W diagramie E-R **klucze częściowe** oznaczane są przez **podkreślenie** etykiety atrybutu **linią przerywaną**.
- Związki między słabą jednostką a jednostkami jej właścicieli nazywamy **związkami identyfikującymi**.
- W notacji graficznej zbiory słabe encji oraz zbiory związków identyfikujących rysujemy **podwójną linią**.



Diagramy E-R → tabele (przykład)



student		
nr_albumu	imie	nazwisko
12345	Jan	Pies
12346	Ewa	Lis
12347	Adam	Słoń

oceny			
nr_albumu	przedmiot	ocena	data
12345	matematyka	2,0	20.01.2006
12346	matematyka	3,0	20.01.2006
12347	wychowanie fizyczne	3,0	20.01.2006
12346	bazy danych	3,0	26.01.2006
12347	bazy danych	5,0	26.01.2006

Diagramy E-R → tabele

- Jeżeli **zbiór związków** nie ma własnego klucza głównego, wówczas tabela musi dodatkowo zawierać odpowiednią do własności kombinację kluczy głównych swoich zbiorów encji.
 - » Np. Związek *wielu do wielu* jest reprezentowany przy pomocy tabeli o dodatkowych kolumnach odpowiadających kluczom głównym zbiorów encji biorących udział w związku.



student		
nr_albumu	imie	nazwisko
12345	Jan	Pies
12346	Ewa	Lis
12347	Adam	Słoń

ma_zajecia	
nr_albumu	ID_przedmiot
12345	1
12345	3
12346	1
12347	1
12346	4
12347	4
12345	4

przedmioty				
ID_przedmiot	rodzaj_studiow	przedmiot	prowadzacy_pesel	forma
1	stacjonarne	matematyka	29035934567	egzamin
2	zaoczne	matematyka	15057698456	egzamin
3	podyplomowe	bazy danych	31169675846	zaliczenie
4	stacjonarne	ekonomia	31169675846	zaliczenie

„Student lubi piwo...”

L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	Alkohol [%]
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Żywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe-Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszeniczne	pszeniczne	5,6
9.	Tadeusz	Konwicky	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

„Student lubi piwo...”

STUDENT			PIWO			
L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	Alkohol [%]
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Żywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe-Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszeniczne	pszeniczne	5,6
9.	Tadeusz	Konwicky	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

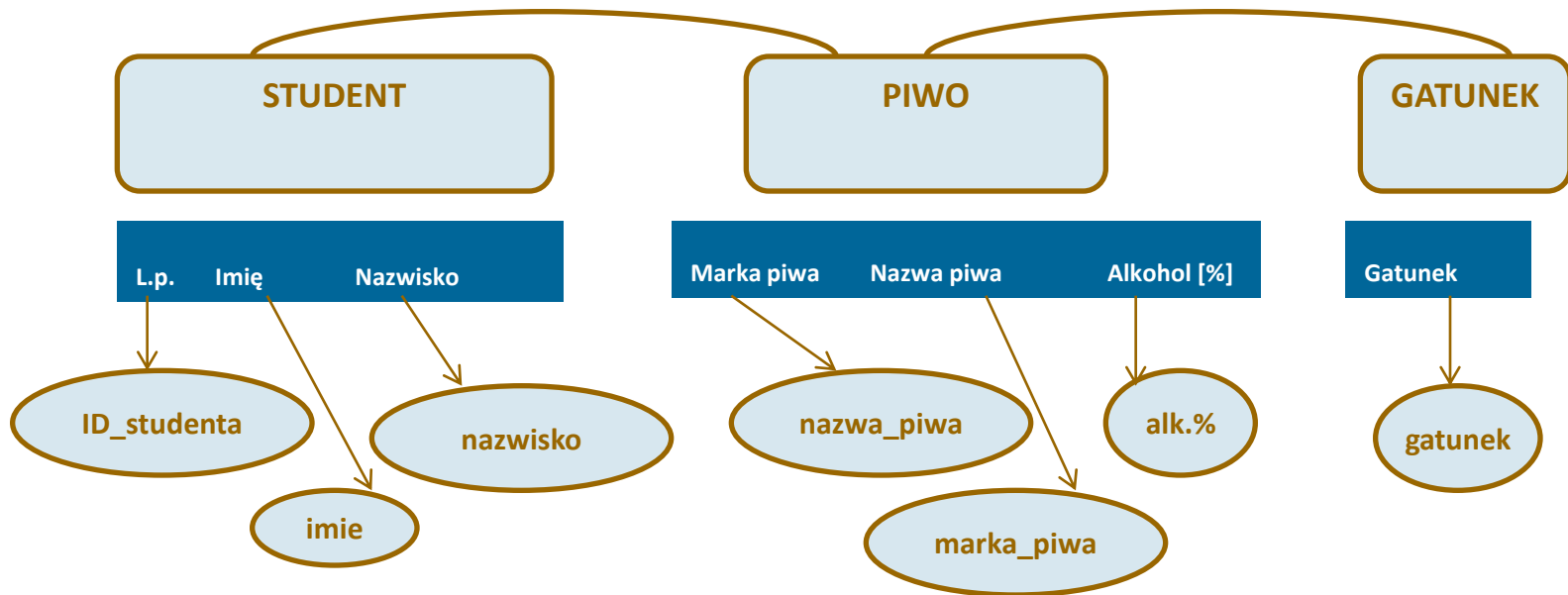
„Student lubi piwo...”

STUDENT			PIWO		GATUNEK	PIWO
L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	Alkohol [%]
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Żywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe-Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszeniczne	pszeniczne	5,6
9.	Tadeusz	Konwicky	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

„Student lubi piwo...”

STUDENT			PIWO		GATUNEK	PIWO
L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	Alkohol [%]
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Żywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszeniczne	pszeniczne	5,6
9.	Tadeusz	Konwicky	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

„Student lubi piwo...”



STUDENT	
ID_studenta	Char(20)
imie	Char(20)
nazwisko	Char(20)

PIWO	
marka_piwa	Char(20)
nazwa_piwa	Char(20)
alk.	Float

GATUNEK	
gatunek	Char(20)


„Student lubi piwo...”


STUDENT	
ID_studenta	Char(20)
imie	Char(20)
nazwisko	Char(20)

PIWO	
marka_piwa	Char(20)
nazwa_piwa	Char(20)
alk.	Float

GATUNEK	
gatunek	Char(20)



STUDENT	
 ID_studenta	Char(20) NN
imie	Char(20)
nazwisko	Char(20)

PIWO	
 ID_piwa	Char(20) NN
marka_piwa	Char(20)
nazwa_piwa	Char(20)
alk.	Float

GATUNEK	
 ID_gatunku	Char(20) NN (PK)
gatunek	Char(20)

„Student lubi piwo...”

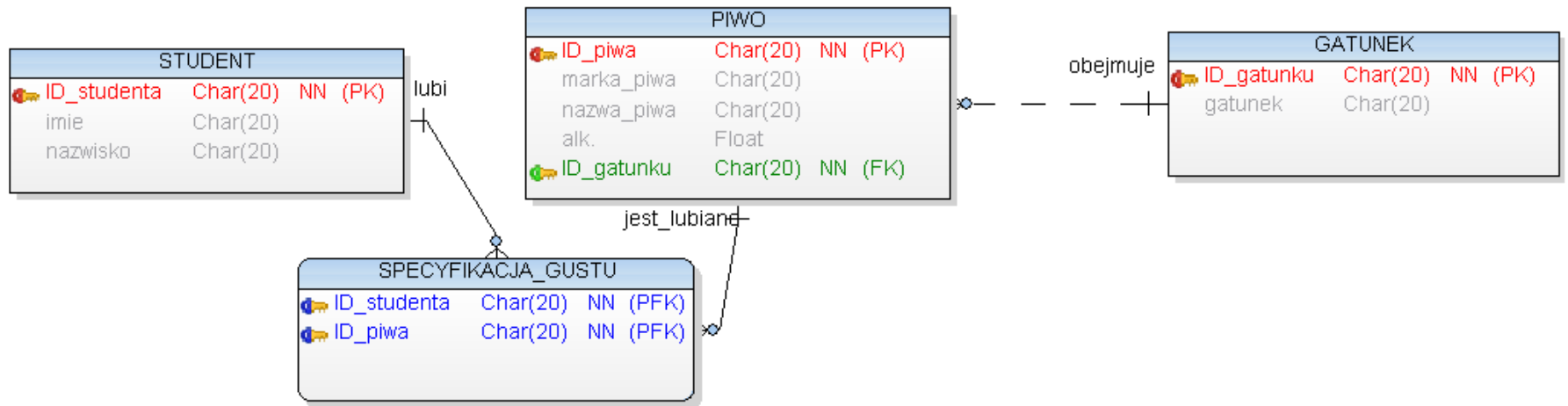
STUDENT			PIWO		GATUNEK	Alkohol [%]
L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Zywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszeniczne	pszeniczne	5,6
9.	Tadeusz	Konwicki	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

„Student lubi piwo...”

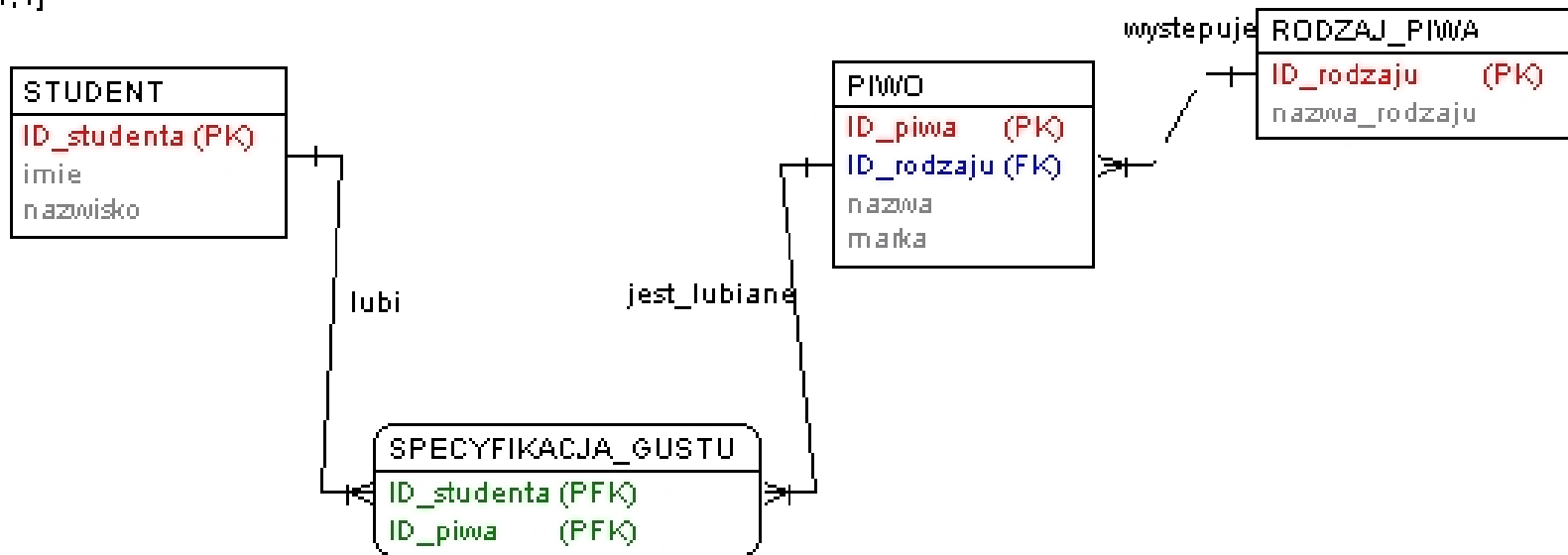
STUDENT	
ID_studenta	Char(20) NN
imie	Char(20)
nazwisko	Char(20)

PIWO	
ID_piwa	Char(20) NN
marka_piwa	Char(20)
nazwa_piwa	Char(20)
alk.	Float

GATUNEK	
ID_gatunku	Char(20) NN (PK)
gatunek	Char(20)



[1.1]



Normalizacja.

Przykład wad bazy danych:

– Rozważmy schemat relacji:

» $pracownik_oddzialu = \langle nazwa_oddzialu, numer_oddzialu, adres_oddzialu, pesel, imie, nazwisko, stanowisko, wynagrodzenie \rangle$

– Redundancja:

- » Dane $nazwa_oddzialu, numer_oddzialu, adres_oddzialu$ są pamiętane dla każdego pracownika
- » Marnowanie miejsca (przestrzeni potrzebnej dla przechowywania danych)
- » Komplikacje (aktualizacja, błąd, usuwanie...)

– Brak możliwości reprezentowania pewnych informacji

- » Nie można reprezentować informacji o oddziałach, których powołanie dopiero jest planowane i nie zatrudniły jeszcze pracowników
- » Rozwiązaniem mogło by być zastosowanie wartości NULL, ale takie rozwiązanie także stwarza pewne problemy
- » Nie możemy zaprezentować w prosty sposób wszystkich pracowników z danego miasta, ponieważ miejscowość nie jest wyodrębnionym atrybutem

Rodzaje anomalii:

- **Anomalia dołączania** - wadą jest to, że musimy wpisywać wszystko albo nic,
 - » np. planuje się powołanie nowego oddziału, nie możemy jednak zapisać w bazie miejscowości, adresu i nazwy zakupionego budynku, dopóki nowy oddział nie zatrudni pracowników
- **Anomalia aktualizacji** - np. oddział firmy został przeniesiony, przez co musimy aktualizować jego adres dla każdego pracownika, a przez przypadek omijamy Jana Kowalskiego
- **Anomalia usuwania** – zamknięto oddział firmy, w związku z czym usunięto z bazy wszystkie rekordy zawierające jego nazwę, tym samym usunięto wszystkie dane dotyczące pracowników

Dekompozycja:

- Można **zdekomponować** schemat relacji – *pracownik_oddzialu*:
 - » *oddzial* = $\langle nazwa_oddzialu, numer_oddzialu, miejscowosc_oddzialu, adres_oddzialu \rangle$
 - » *pracownik* = $\langle numer_oddzialu, pesel, imie, nazwisko, stanowisko, wynagrodzenie \rangle$
- **Wszystkie atrybuty** oryginalnego schematu (R) muszą się pojawić w dekompozycji (R_1, R_2):

$$R = R_1 \cup R_2$$

- W przypadku gdy relacja nie posiada właściwej postaci należy dokonać dekompozycji relacji R na $\langle R_1, R_2, \dots, R_n \rangle$
- Proces „dochodzenia” do „właściwej” postaci określa się mianem **normalizacji**

Postacie normalne:

- Pierwsza (1PN)
- Druga (2PN)
- Trzecia (3PN)
- Boyce'a-Codd'a (PNBC)
- Czwarta (4PN)
- Piąta (5PN)

Pierwsza postać normalna

- Relacja jest w pierwszej postaci normalnej, jeśli każda wartość atrybutu w każdej krotce tej relacji jest **wartością elementarną**, czyli nierozkładalną.
- Relacja jest w pierwszej postaci normalnej, jeśli nie ma **powtarzających się grup**.

Przykład:

pracownik oddzialu								
numer_ oddzialu	nazwa_ o ddzialu	miestowosc_ oddzialu	adres_ oddzialu	pesel	imie	nazwisko	stanowisko	wynagrodze nie
A1	Alfaton	Kraków	Słoneczna 2/5	21108106713	Jan	Głuchomolski	Sprzedawca	2 700,00 zł
A1	Alfaton	Kraków	Słoneczna 2/5	15057698456	Wiktor	Sakowski	Kierowca	1 800,00 zł
A2	Alfabras	Kraków	Fabryczna 13/7	13015778123	Magdalena	Waligórska	Sprzedawca	2 700,00 zł
B1	Betatrix	Hel	Leśna 10	30127623435	Ignacy	Kuchta	Kierownik sekcji	3 500,00 zł
B1	Betatrix	Hel	Leśna 10	29035934567	Marzena	Sępieł	Kierowca	1 800,00 zł
B1	Betatrix	Hel	Leśna 10	31169675846	Jadwiga	Makusz	Sprzedawca	2 700,00 zł
B2	Betanor	Gdynia	Widokowa 12/26	21016922987	Stanisław	Marekowska	Kierowca	1 800,00 zł
B3	Betarex	Gdynia	Portowa 23	17037487678	Zuzanna	Wał	Sprzedawca	2 700,00 zł

oddzial			
numer_ oddzialu	nazwa_ oddzialu	miestowosc_ oddzialu	adres_ oddzialu
A1	Alfaton	Kraków	Słoneczna 2/5
A2	Alfabras	Kraków	Fabryczna 13/7
B1	Betatrix	Hel	Leśna 10
B2	Betanor	Gdynia	Widokowa 12/26
B3	Betarex	Gdynia	Portowa 23

pracownik					
pesel	imie	nazwisko	stanowisko	wynagrodzenie	numer_ oddzialu
21108106713	Jan	Głuchomolski	Sprzedawca	2 700,00 zł	A1
15057698456	Wiktor	Sakowski	Kierowca	1 800,00 zł	A1
13015778123	Magdalena	Waligórska	Sprzedawca	2 700,00 zł	A2
30127623435	Ignacy	Kuchta	Kierownik sekcji	3 500,00 zł	B1
29035934567	Marzena	Sępieł	Kierowca	1 800,00 zł	B1
31169675846	Jadwiga	Makusz	Sprzedawca	2 700,00 zł	B1
21016922987	Stanisław	Marekowska	Kierowca	1 800,00 zł	B2
17037487678	Zuzanna	Wał	Sprzedawca	2 700,00 zł	B3

Druga i trzecia postać normalna:

- Relacja jest w drugiej postaci normalnej, jeśli jest w 1PN oraz każdy atrybut tej relacji **nie wchodzący** w skład żadnego klucza potencjalnego jest w **pełni funkcyjnie zależny** od wszystkich kluczy potencjalnych tej relacji.
- Relacja jest w 2PN jeżeli każdy atrybut nie wchodzący w skład klucza **zależy od klucza a nie od jego części**.
- Relacja będąca w pierwszej postaci normalnej, jest równocześnie w drugiej postaci normalnej, jeśli wszystkie jej klucze potencjalne są **kluczami prostymi**.
- Dana relacja jest w trzeciej postaci normalnej, jeśli jest ona w drugiej postaci normalnej i każdy jej atrybut nie wchodzący w skład żadnego klucza potencjalnego **nie jest przechodnio funkcyjnie zależny** od żadnego klucza potencjalnego tej relacji.

Druga i trzecia postać normalna:

Inaczej mówiąc, **wszystkie niekluczowe kolumny są określane**

**kluczem, całym kluczem i tylko kluczem,
„...tak nam dopomóż Codd”**

Przykład:

Protokoły zaliczeń i egzaminów

Rodzaj studiów	Rok studiów	Rok akademicki	Przedmiot	Prowadzący	Forma	Termin	Student		Cena	Data
				Imię i nazwisko			nr albumu	Nazwisko i imię		
stacjonarne	I	2005/2006	matematyka	prof. Jan Kot	egzamin	I	12345	Pies Jan	2,0	20.01.2006
stacjonarne	I	2005/2006	matematyka	prof. Jan Kot	egzamin	I	12346	Lis Ewa	3,0	20.01.2006
zaoczne	I	2005/2006	matematyka	prof. Jan Kot	egzamin	II	12347	Stoń Adam	3,0	20.01.2006
stacjonarne	I	2005/2006	wychowanie fizyczne	mgr Janina Mysz	zaliczenie	I	12345	Pies Jan	5,0	26.01.2006
stacjonarne	I	2005/2006	wychowanie fizyczne	mgr Janina Mysz	zaliczenie	I	12346	Lis Ewa	3,0	26.01.2006
zaoczne	I	2005/2006	wychowanie fizyczne	mgr Janina Mysz	zaliczenie	II	12347	Stoń Adam	3,0	26.01.2006

Wartości atrybutów nie są elementarne

Przykład (c.d.):

Klucz potencjalny: <rodzaj_studiow, rok_studiow, rok_akademicki, przedmiot, forma, termin, student_nr_albumu>

Zależności funkcyjne

Protokoły zaliczeń i egzaminów

rodzaj_studiow	rok_studiow	rok_akademicki	przedmiot	przewadzacy_tytul	przewadzacy_imie	przewadzacy_nazwisko	forma	termin	student_nr_albumu	student_imie	student_nazwi	ocena	data
stacjonarne	I	2005/2006	matematyka	prof.	Jan	Kot	egzamin	I	12345	Jan	Pies	2,0	20.01.2006
stacjonarne	I	2005/2006	matematyka	prof.	Jan	Kot	egzamin	I	12346	Ewa	Lis	3,0	20.01.2006
zaoczne	I	2005/2006	matematyka	prof.	Jan	Kot	egzamin	II	12347	Adam	Stoń	3,0	20.01.2006
stacjonarne	I	2005/2006	wychowanie fizyczne	mgr	Janina	Mysz	zaliczenie	I	12345	Jan	Pies	5,0	26.01.2006
stacjonarne	I	2005/2006	wychowanie fizyczne	mgr	Janina	Mysz	zaliczenie	I	12346	Ewa	Lis	3,0	26.01.2006
zaoczne	I	2005/2006	wychowanie fizyczne	mgr	Janina	Mysz	zaliczenie	II	12347	Adam	Stoń	3,0	26.01.2006

Powtarzające się grupy

prowadzacy			
pesel	tytul	imie	nazwisko
29035934567	prof.	Jan	Kot
31169675846	mgr	Janina	Mysz
15057698456	dr hab. inż.	Wojciech	Kuna

przedmioty						
id_przedmiot	rodzaj_studiow	rok_studiow	rok_akademicki	przedmiot	prowadzacy_pesel	forma
1	stacjonarne	I	2005/2006	matematyka	29035934567	egzamin
2	zaoczne	II	2005/2006	matematyka	15057698456	egzamin
3	stacjonarne	I	2005/2006	wychowanie fizyczne	31169675846	zaliczenie
4	zaoczne	II	2005/2006	wychowanie fizyczne	31169675846	zaliczenie

protokoly				
id_przedmiot	termin	nr_albumu	ocena	data
1	I	12345	2,0	20.01.2006
1	I	12346	3,0	20.01.2006
2	II	12347	3,0	20.01.2006
3	I	12345	5,0	26.01.2006
3	I	12346	3,0	26.01.2006
4	I	12347	3,0	26.01.2006

student		
nr_albumu	imie	nazwisko
12345	Jan	Pies
12346	Ewa	Lis
12347	Adam	Stoń

Klucz główny relacji protokoly:
<id_przedmiot, termin, nr_albumu>

Ta relacja nie jest w trzeciej postaci normalnej

Pracownik	PESEL	KodPocztowy	Miejscowość	Województwo
Jan Kowalski	12345678911	32-082	Bolechowice	małopolskie
Adam Kot	98977796666	30-150	Kraków	małopolskie
Ewa Lis	76281976372	32-082	Bolechowice	małopolskie



Podsumowanie:

- Normalizacja ma na celu takie przekształcenie relacji, by uniknąć **redundancji i anomalii**.
- Przekształcenie relacji do kolejnych postaci normalnych wiąże się najczęściej ze **zmniejszeniem ilości pamięci** potrzebnej do **przechowania informacji**.
- Unikanie powtórzeń pozwala na **łatwiejszą i szybszą aktualizację** danych.
- Doprowadzenie bazy do wysokiej postaci normalizacji może **spowolnić odczyt** w dużych bazach ze względu na skomplikowany schemat danych.
- W większości przypadków po znormalizowaniu bazy danych przychodzi kolej na rozważenie możliwości wykonania **odwrotnej operacji (denormalizacji)**, polegającej na połączeniu niektórych znormalizowanych tabel, a to z myślą o przyspieszeniu dostępu do pewnych danych.

