



Wykład 3

# Normalizacja



# 6. Normalizacja

## ★ TEMATYKA:

- Cel normalizacji
- Etapy normalizacji
- Związki zależności funkcyjne i nefunkcyjne
- Postacie normalne
- Diagramy zależności
- Akomodacja – przekształcanie diagramu zależności w schemat relacyjny

# Opis modelu relacyjnego w notacji nawiasowej

- ★ Metoda pozwala w postaci skróconej opisać definicję schematu bazy.
- ★ W opisie najpierw podajemy nazwę relacji (tabeli) a następnie w nawiasie kolejne atrybuty. Jako pierwsze podaje się atrybuty stanowiące klucz główny jednocześnie podkreślając je.

**Dla przykładu opisanego wcześniej zapis będzie miał postać:**

Moduły(NazwaModułu, NrPrac)

Wykładowcy(NrPrac, Pracownik)

Oceny(NazwaModułu, NrStud, TypOceny, Ocena)

Studenci(NrStudenta, Student)

## 6.1. Cel normalizacji

- ★ **Normalizacja to proces upraszczania struktury bazy danych w taki sposób, aby osiągnęła ona postać optymalną.**
- ★ Normalizacja wykonuje się na etapie projektowania modelu fizycznego danych
- ★ **Dzięki normalizacji można uniknąć anomalii – błędów lub niespójności w bazie danych (w tym również redundancji).**
- ★ **Można wyróżnić 3 rodzaje anomalii:**
  - anomalie przy wstawianiu rekordu
    - dopisanie rekordu powoduje dezaktualizację innego pola,
  - anomalie przy usuwaniu rekordu
    - usunięcie wiersza powoduje usunięcie większej ilości informacji niż żeśmy zamierzali,
  - anomalie przy modyfikacji rekordu
    - zmiana jednego rekordu powoduje konieczność zmiany zapisów w innych rekordach.

# Przykład

- ★ Dysponujemy bazą danych z informacjami o studentach, modułach oraz wykładowcach na uniwersytecie

| Moduły                               |        |                |        |             |       |          |
|--------------------------------------|--------|----------------|--------|-------------|-------|----------|
| NazwaModułu                          | NrPrac | Pracownik      | NrStud | Student     | Ocena | TypOceny |
| Systemy relacyjnych baz danych       | 244    | Buczek Jan     | 34698  | Kowalski H. | 4.0   | zal.     |
| Systemy relacyjnych baz danych       | 244    | Buczek Jan     | 34698  | Kowalski H. | 3.5   | egz.     |
| Systemy relacyjnych baz danych       | 244    | Buczek Jan     | 37653  | Nowak R.    | 3.0   | zal.     |
| Systemy relacyjnych baz danych       | 244    | Buczek Jan     | 34610  | Lech M.     | 5.0   | zal.     |
| Projektowanie relacyjnych baz danych | 244    | Buczek Jan     | 34698  | Kowalski H. | 3.0   | zal.     |
| Projektowanie relacyjnych baz danych | 244    | Buczek Jan     | 34698  | Kowalski H. | 4.0   | egz.     |
| Obiektowe bazy danych                | 445    | Kalita Henryk  | 35785  | Woś S.      | 3.5   | egz.     |
| Roproszone bazy danych               | 247    | Wysocki Edward | 34789  | Janda K.    | 5.0   | zal.     |

- Co będzie gdy usuniemy studenta **Wosia**? – **stracimy informację o Obiektowych bazach danych i wykładowcy Kalicie Henryku** – **anomalia przy usuwaniu**
- Co będzie gdy zmienimy wykładowcę Rozproszonych baz danych? – **musimy zmodyfikować dwa pola: NrPrac oraz Pracownik** – **anomalia przy modyfikacji**
- Co będzie gdy wpiszemy nowego studenta na moduł? – **możliwe to będzie dopiero po uzyskaniu przez niego pierwszego zaliczenia** – **anomalia przy wstawianiu**

## 6.2. Etapy normalizacji

1. Zebranie danych
2. Przekształcenie do pierwszej postaci normalnej (1PN)
3. Przekształcenie do drugiej postaci normalnej (2PN)
4. Przekształcenie do trzeciej postaci normalnej (3PN)

Po znormalizowaniu do 3 PN najczęściej tablice są już pozbawione **anomalii**, jeżeli nie to należy je:

- ✳ Przekształcić do postaci normalnej Boyce'a-Codd'a (BCNF)
- ✳ Przekształcić do czwartej postaci normalnej (4PN)
- ✳ Przekształcić do piątej postaci normalnej (5PN)

Proces normalizacji jest włożony. To znaczy, że każda wyższa postać normalna jest podzbiorem postaci niższej.

## 6.2.1 Rozkład relacji (tablic) a normalizacja

**Proces przekształcania nieznormalizowanego zbioru danych w pełni znormalizowaną bazę danych nosi nazwę dekompozycji odwracalnej (rozkładu odwracalnego)**

### **★ Cechy dekompozycji odwracalnej**

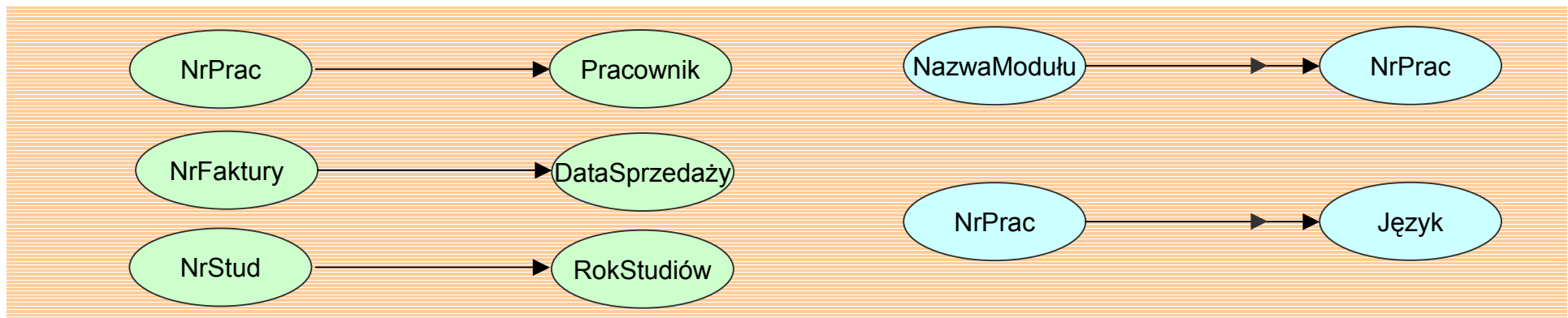
- usuwa redundancję z relacji
- można ją odwrócić przez naturalne złączenie
- powinna doprowadzić relacje do tzw. postaci normalnej
- nie powinna powodować utraty zależności istniejących w relacji pierwotnej

## 6.3. Zależności funkcyjne i nefunkcyjne (wielowartościowe)

- ★ Dwa elementy A i B są w związku zależności (związku determinowania), jeżeli pewne wartości elementu danych B występują zawsze z pewnymi wartościami elementu A.
- ★ Zależność funkcyjna (determinowanie) między elementami danych wskazuje kierunek związku. Jeżeli A determinuje B to związek jest od A do B nie odwrotnie

Element danych B jest **funkcyjnie zależny** od elementu danych A, jeżeli dla każdej wartości A istnieje jedna jednoznacznie określona wartość B

Element danych B jest **niefunkcyjnie zależny** od elementu danych A, jeżeli dla każdej wartości elementu danych A istnieje ograniczony zbiór wartości elementu B



- ★ 1 PN, 2 PN, 3 PN i BCNF dotyczą zależności funkcyjnych
- ★ 4 PN oraz 5 PN dotyczą zależności nefunkcyjnych



## 6.3.1. Zależności funkcyjne

- ★ Najważniejszy rodzaj więzów, z jakimi mamy doczynienia w modelu relacyjnym, dotyczy więzów jednoznaczności, które nazywa się również **zależnością funkcyjną**.
- ★ Nie istnieją metody pozwalające automatycznie określić zależności funkcyjne, aby to uczynić należy dokładnie przeanalizować znaczenie wszystkich atrybutów.

| Filmy           |      |         |       |           |                |
|-----------------|------|---------|-------|-----------|----------------|
| Tytuł           | Rok  | Długość | Typ   | Producent | Gwiazda        |
| Gwiezdne Wojny  | 1977 | 124     | kolor | Fox       | Carrie Fisher  |
| Gwiezdne Wojny  | 1977 | 124     | kolor | Fox       | Mark Hamil     |
| Gwiezdne Wojny  | 1977 | 124     | kolor | Fox       | Harrison Ford  |
| Potężne Kaczory | 1991 | 104     | kolor | Disney    | Emilio Estavez |

W relacji można wyodrębnić zależności:

tytuł, rok → długość

tytuł, rok → typ

tytuł, rok → producent

Ponieważ wszystkie zależności mają po lewej stronie te same atrybuty można je zapisać:

tytuł, rok → długość, typ, producent

- ★ Zależności funkcyjne (podobnie jak inne więzy) dotyczą schematu bazy danych, a nie określonej instancji

# Reguły dotyczące zależności funkcyjnych

★ Reguła przechodności

# Aksjomaty Armstronga

- ★ Zbiór atrybutów określa funkcyjnie dowolny jego podzbiór (Zwrotność)

- np.:  $\text{nr}, \text{nazwisko} \rightarrow \text{nazwisko}$

- ★ Jeżeli zbiór atrybutów  $X$  funkcyjnie określa zbiór atrybutów  $Y$  oraz  $Z$  jest innym zbiorem atrybutów wówczas suma zbiorów  $X$  i  $Z$  funkcyjnie określa sumę  $Y$  i  $Z$

$$X \rightarrow Y \Rightarrow X \cup Z \rightarrow Y \cup Z$$

(Rozszerzenie)

- np.:  $\text{nr} \rightarrow \text{nazwisko}$  to zachodzi również  $\text{nr}, \text{data} \rightarrow \text{nazwisko}, \text{data}$

- ★ Jeżeli  $X \rightarrow Y$  i  $Y \rightarrow Z \Rightarrow X \rightarrow Z$

(Przechodniość)

- $X \rightarrow Y \Rightarrow X \cup Z \rightarrow Z$

- $X \rightarrow Y$  i  $X \rightarrow Z \Leftrightarrow X \rightarrow Y \cup Z$

- Jeżeli  $A = \{A_1, A_2 \dots A_n\}$  to  $X \rightarrow A \Leftrightarrow X \rightarrow A_1 \dots X \rightarrow A_n$

# Obliczanie domknięcia zbioru atrybutów

★ Bardzo często należy określić, które pojedyncze atrybuty są funkcyjnie zależne od danego zbioru atrybutów.

★ Algorytm X-domkniętości:

Krok1:  $X(0) = X$ ,  $n=0$

Krok2: Jeżeli istnieje zależność  $A \rightarrow B$  oraz  $A \subset X(n)$  i  $B \not\subset X(n)$  to  $X(n+1) = X(n) \cup B$ .

W przeciwnym wypadku zakończ algorytm

Krok 3:  $n=n+1$  i wróć do kroku 2.

Rozważmy relację

Należności(nazwisko, ulica, miasto, województwo, data, wielkość)

Z następującymi zależnościami funkcyjnymi:

$\text{nazwisko} \rightarrow \text{ulica, miasto, województwo}$

$\text{nazwisko, data} \rightarrow \text{wielkość}$

$\text{miasto} \rightarrow \text{województwo}$

Wykonując algorytm X-domkniętości otrzymamy:

$X(0) = \{\text{nazwisko, data}\}$

$X(1) = \{\text{nazwisko, ulica, miasto, województwo, data}\}$

$X(2) = \{\text{nazwisko, ulica, miasto, województwo, data, wielkość}\}$

$X(3) = X(2)$

## 6.4. Pierwsza postać normalna 1PN

**Relacja (tablica) jest w pierwszej postaci normalnej (1PN) wtedy i tylko wtedy, gdy każdy atrybut niekluczowy jest funkcjonalnie zależny od klucza głównego**

- ✳ Pierwsza postać normalna to warunek, że wszystkie wartości kolumn muszą być elementarne
- ✳ Elementarne znaczy w tym przypadku niepodzielne. 1 PN wymaga, żeby dla każdej pozycji wiersz-kolumna w tablicy istniała tylko jedna wartość, a nie tablica lub lista wartości.
- ✳ Jeśli w kolumnie przechowuje się całe listy wartości, wtedy trudno jest nimi operować.
- ✳ 1 PN zabrania także istnienia powtarzających się grup, nawet jeśli miałyby być one złożone z kolumn elementarnych

# Przejdźcie do 1 PN – przykład

| Moduły |                                      |        |               |        |             |       |          |
|--------|--------------------------------------|--------|---------------|--------|-------------|-------|----------|
| KP     | NazwaModułu                          | NrPrac | Pracownik     | NrStud | Student     | Ocena | TypOceny |
|        | Systemy relacyjnych baz danych       | 244    | Buczek Jan    | 34698  | Kowalski H. | 4.0   | zal.     |
|        |                                      |        |               |        |             | 3.5   | egz.     |
|        |                                      |        |               | 34610  | Lech M.     | 5.0   | zal.     |
|        | Projektowanie relacyjnych baz danych | 244    | Buczek Jan    | 34698  | Kowalski H. | 3.0   | zal.     |
|        |                                      |        |               |        |             | 4.0   | egz.     |
|        | Obiektowe bazy danych                | 445    | Kalita Henryk | 35785  | Woś S.      | 3.5   | egz.     |

★ Atrybuty NrStud, Student, Ocena oraz TypOceny nie są funkcyjnie zależne od klucza głównego, pozostałe są zależne. Należy więc utworzyć dwie tabele:

- jedna dla atrybutów zależnych
- drugą dla funkcyjnie niezależnych atrybutów

| Moduły      |        |           |
|-------------|--------|-----------|
| NazwaModułu | NrPrac | Pracownik |

| Oceny |                                      |        |          |             |       |
|-------|--------------------------------------|--------|----------|-------------|-------|
| KP    | NazwaModułu                          | NrStud | TypOceny | Student     | Ocena |
|       | Systemy relacyjnych baz danych       | 34698  | zal.     | Kowalski H. | 4.0   |
|       | Systemy relacyjnych baz danych       | 34698  | egz.     | Kowalski H. | 3.5   |
|       | Systemy relacyjnych baz danych       | 34610  | zal.     | Lech M.     | 5.0   |
|       | Projektowanie relacyjnych baz danych | 34698  | zal.     | Kowalski H. | 3.0   |
|       | Projektowanie relacyjnych baz danych | 34698  | egz.     | Kowalski H. | 4.0   |
|       | Obiektowe bazy danych                | 35785  | egz.     | Woś S.      | 3.5   |

Atrybuty NazwaModułu, NrStud i TypOceny utworzą klucz główny tabeli Oceny

## 6.5. Druga postać normalna 2PN

**Relacja jest w drugiej postaci normalnej wtedy i tylko wtedy, gdy jest w 1PN i każdy atrybut niekluczowy jest w pełni funkcyjnie zależny od klucza głównego**

- ★ Tablica ma drugą postać normalną, jeśli jest w 1 PN i każda kolumna nie należąca do żadnego klucza potencjalnego jest całkowicie zależna od klucza głównego.
- ★ Innymi słowy, tablice powinny przechowywać dane dotyczące tylko jednej „rzeczy” (jednostki, obiektu, zdarzenia) oraz ta „rzecz” powinna być opisywalna przez jej klucz główny.
- ★ Tablica która jest w 1PN może nie być w 2PN tylko wtedy gdy posiada klucz główny złożony bo tylko wówczas któryś z atrybutów może być identyfikowany przez część klucza.

# Przejsie do 2 PN – przykład

| Oceny |                                      |        |          |             |       |
|-------|--------------------------------------|--------|----------|-------------|-------|
| KP    | NazwaModułu                          | NrStud | TypOceny | Student     | Ocena |
|       | Systemy relacyjnych baz danych       | 34698  | zal.     | Kowalski H. | 4.0   |
|       | Systemy relacyjnych baz danych       | 34698  | egz.     | Kowalski H. | 3.5   |
|       | Systemy relacyjnych baz danych       | 34610  | zal.     | Lech M.     | 5.0   |
|       | Projektowanie relacyjnych baz danych | 34698  | zal.     | Kowalski H. | 3.0   |
|       | Projektowanie relacyjnych baz danych | 34698  | egz.     | Kowalski H. | 4.0   |
|       | Obiektowe bazy danych                | 35785  | egz.     | Woś S.      | 3.5   |

Tabela Oceny (niebieska) nie jest w 2PN gdyż pole Student nie zależy od całego klucza głównego a jedynie od atrybutu NrStud

Tabele w 2PN

| Oceny |                                      |        |          |       |
|-------|--------------------------------------|--------|----------|-------|
| KP    | NazwaModułu                          | NrStud | TypOceny | Ocena |
|       | Systemy relacyjnych baz danych       | 34698  | zal.     | 4.0   |
|       | Systemy relacyjnych baz danych       | 34698  | egz.     | 3.5   |
|       | Systemy relacyjnych baz danych       | 34610  | zal.     | 5.0   |
|       | Projektowanie relacyjnych baz danych | 34698  | zal.     | 3.0   |
|       | Projektowanie relacyjnych baz danych | 34698  | egz.     | 4.0   |
|       | Obiektowe bazy danych                | 35785  | egz.     | 3.5   |

| Studenci |             |
|----------|-------------|
| NrStud   | Student     |
| 34698    | Kowalski H. |
| 34610    | Lech M.     |
| 35785    | Woś S.      |

| Moduły |                                      |        |               |
|--------|--------------------------------------|--------|---------------|
| KP     | NazwaModułu                          | NrPrac | Pracownik     |
|        | Systemy relacyjnych baz danych       | 244    | Buczek Jan    |
|        | Projektowanie relacyjnych baz danych | 244    | Buczek Jan    |
|        | Obiektowe bazy danych                | 445    | Kalita Henryk |



## 6.6. Trzecia postać normalna 3PN

**Relacja jest w trzeciej postaci normalnej wtedy i tylko wtedy, gdy jest w 2PN i każdy atrybut niekluczowy jest bezpośrednio zależny od klucza głównego.**

- ★ Tablica jest w trzeciej postaci normalnej jeśli jest w 2 PN i wszystkie kolumny nie należące do żadnego klucza potencjalnego są wzajemnie niezależne.**

# Przejdźcie do 3 PN – przykład

| Moduły                               |        |               |
|--------------------------------------|--------|---------------|
| NazwaModułu                          | NrPrac | Pracownik     |
| Systemy relacyjnych baz danych       | 244    | Buczek Jan    |
| Projektowanie relacyjnych baz danych | 244    | Buczek Jan    |
| Obiektowe bazy danych                | 445    | Kalita Henryk |

Jedynie tabela *Moduły* nie jest w 3PN gdyż pole *Pracownik* zależy od pola *NrPrac* a nie bezpośrednio od pola *NazwaModułu*.

Należy dokonać podziału tabeli

Tabele w 3 PN

| Moduły                               |        |
|--------------------------------------|--------|
| NazwaModułu                          | NrPrac |
| Systemy relacyjnych baz danych       | 244    |
| Projektowanie relacyjnych baz danych | 244    |
| Obiektowe bazy danych                | 445    |

| Wykładowcy |               |
|------------|---------------|
| NrPrac     | Pracownik     |
| 244        | Buczek Jan    |
| 445        | Kalita Henryk |

| Oceny |                                      |        |          |       |
|-------|--------------------------------------|--------|----------|-------|
| KP    | NazwaModułu                          | NrStud | TypOceny | Ocena |
|       | Systemy relacyjnych baz danych       | 34698  | zal.     | 4.0   |
|       | Systemy relacyjnych baz danych       | 34698  | egz.     | 3.5   |
|       | Systemy relacyjnych baz danych       | 34610  | zal.     | 5.0   |
|       | Projektowanie relacyjnych baz danych | 34698  | zal.     | 3.0   |
|       | Projektowanie relacyjnych baz danych | 34698  | egz.     | 4.0   |
|       | Obiektowe bazy danych                | 35785  | egz.     | 3.5   |

| Studenci |             |
|----------|-------------|
| NrStud   | Student     |
| 34698    | Kowalski H. |
| 34610    | Lech M.     |
| 35785    | Woś S.      |

# Metoda Bersteina

★ Dowolną relację można przekształcić do 3PN korzystając z metody Bernsteina:

Krok 1. Przekształć każdą zależność tak aby po prawej stronie był tylko jeden atrybut.

Krok 2. Wyeliminuj powtarzające się zależności

Krok 3. Przekształć zależność tak aby żaden podzbiór atrybutów stojący po lewej stronie nie określał prawej strony zależności.

Krok 4. Połącz zależności z takimi samymi lewymi stronami

Krok 5. Znajdź klucz dla relacji pierwotnej. Jeżeli żaden z kluczy nie jest zawarty w zbiorach z kroku poprzedniego to utwórz nowy zbiór z atrybutami klucza.

Krok 6. Jeżeli jakiś nowo utworzony zbiór jest rzutem (projekcją) innego, to wyeliminuj ten zbiór

★ Rozważmy relację

Magazyn(nr, mistrz, wydział, materiał, ilość, cena)

Z następującymi zależnościami funkcyjnymi:

nr → mistrz

mistrz → wydział

mistrz, materiał → ilość

materiał → cena

Krok 1. Zależności funkcyjne spełniają wymagania

Krok 2. Brak zależności redundacyjnych

Krok 3. Lewe strony są minimalne. Żaden podzbiór atrybutów z lewej strony nie określa prawej strony

Krok 4,5 . Otrzymujemy pięć relacji

R1(nr, mistrz)

R2(mistrz, wydział)

R3(mistrz, materiał, ilość)

R4(materiał, cena)

R5(nr, materiał)

Krok 6. Żadna z otrzymanych relacji nie jest rzutem innej \*nie zawiera się w niej)

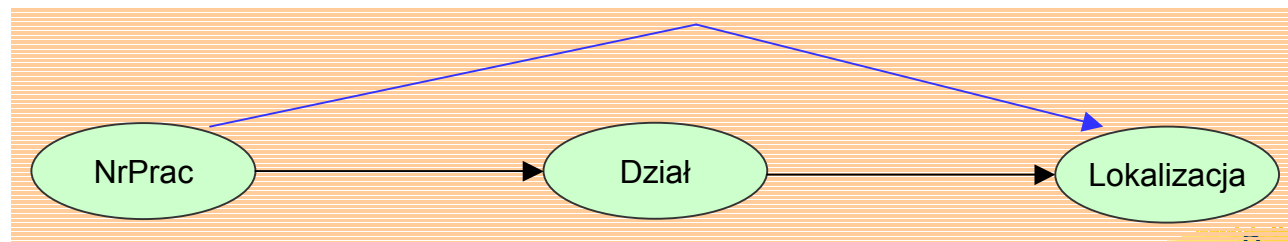
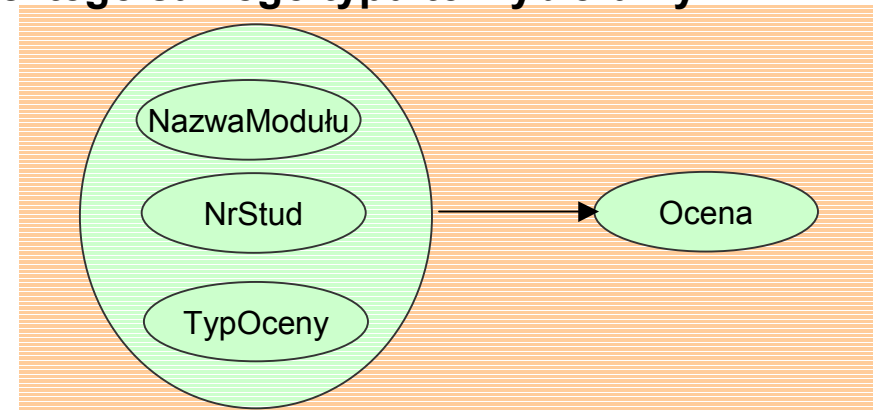
| Magazyn |            |         |          |       |      |
|---------|------------|---------|----------|-------|------|
| Nr      | Mistrz     | Wydział | Materiał | Ilość | Cena |
| 1       | Malinowski | 1       | deski    | 3     | 80   |
| 2       | Kowalski   | 1       | parkiet  | 20    | 95   |
| 3       | Malinowski | 1       | panele   | 50    | 35   |
| 4       | Nowak      | 2       | drzwi    | 1     | 850  |
| 5       | Kowalski   | 1       | drzwi    | 2     | 850  |

## 6.7. Diagramy zależności

- ★ **Klasyczna normalizacja** opisana jako proces rozkładu odwracalnego ma kilka wad:
  - wymaga aby zbiór danych był w pełni określony,
  - jest bardzo czasochłonna.
- ★ Alternatywą dla klasycznej normalizacji mogą być **diagramy zależności**.
- ★ **Zalety diagramów zależności:**
  - nie wymagają pełnego określenia danych
  - czytelny zapis graficzny

## 6.7.2. Pragmatyka rysowania diagramów zależności

- ★ Zależność między dwoma elementami danych można rysować tylko w jedną stronę (od A do B albo od B do A).
- ★ Jeżeli między elementami istnieje w jedną stronę zależność funkcyjna a w drugą niefunkcyjna to wybieramy kierunek zależności funkcyjnej.
  - Np.: w kierunku od `NrPrac` do `NazwaWydziału` jest zależność funkcyjna a w kierunku `NazwaWydziału` do `NrPrac` niefunkcyjna.
- ★ Jeżeli w obu kierunkach występują zależności tego samego typu to wybieramy kierunek, który dla nas jest wygodniejszy.
- ★ Zależności mogą być złożone – gdy złożenie kilku elementów determinuje jakiś inny element
- ★ Jeżeli A determinuje B a B determinuje C to mamy do czynienia z zależnością przechodnią. Wykrycie i usunięcie tego typu zależności jest ważnym elementem procesu normalizacji

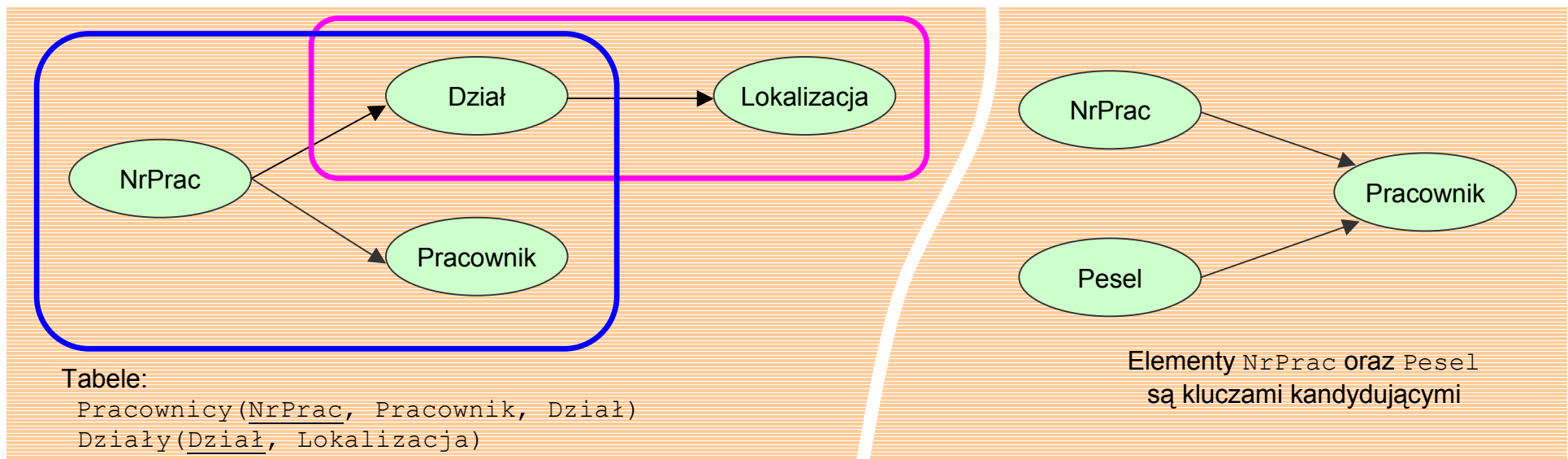


## 6.8 Akomodacja.

**Akomodacja to proces przekształcania diagramu zależności w schemat relacyjny**

### Reguła Boyce'a-Codda

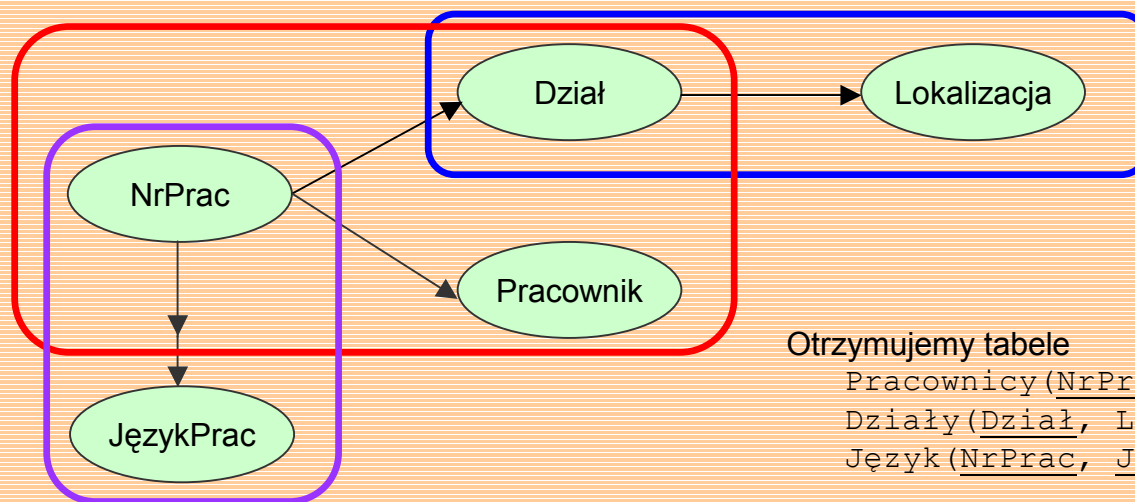
Każdy funkcyjnie determinujący element staje się kluczem kandydującym tabeli. Wszystkie bezpośrednio zależne od niego elementy danych stają się niegłównymi atrybutami tabeli



- ★ Liczba elementów determinujących (z których wychodzą strzałki) wskazuje liczbę wymaganych tabel
- ★ Element do którego wchodzi strzałka i z którego wychodzi strzałka stanowi klucz obcy

Każdy niefunkcyjnie determinujący element staje się częścią klucza głównego tabeli.

- ★ Dokładnie tworzymy klucz główny z determinującego elementu danych i zależnych elementów danych wchodzących w skład związku niefunkcyjnego



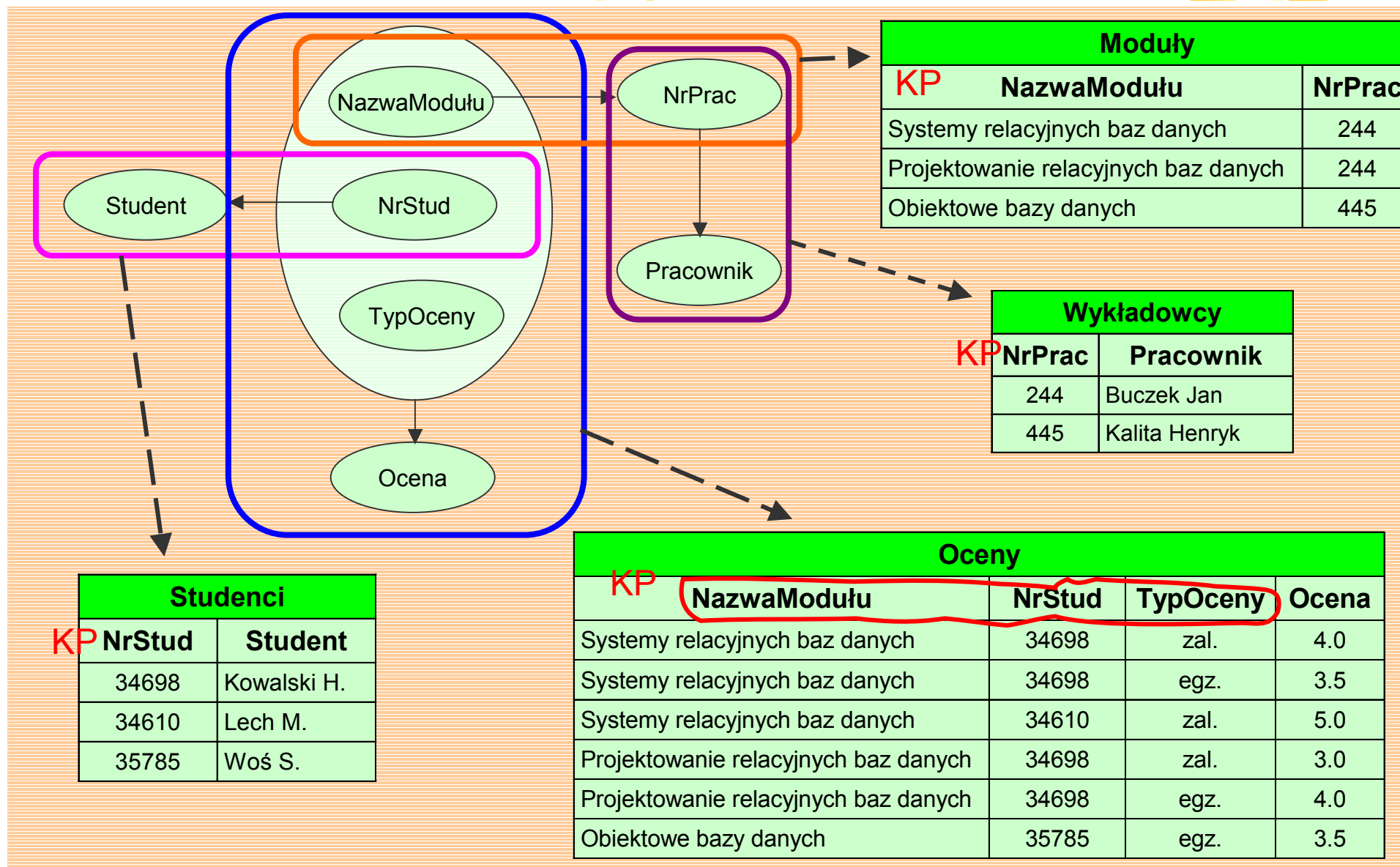
Otrzymujemy tabele

Pracownicy(NrPrac, Pracownik, Dział ...)

Działy(Dział, Lokalizacja ...)

Język(NrPrac, JęzykPrac ...)

## 6.9. Rysowanie diagramów zależności i postacie normalne





## 6.10. Postać normalna Boyce'a-Codda (BCNF)

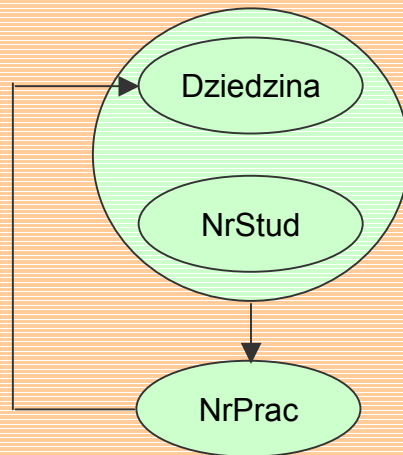
Relacja jest w postaci normalnej Boyce'a-Codda, jeśli tylko z nietrywialnych zależności wynika, że pewien nadklucz wyznacza funkcyjnie jakiś inny atrybut.

- ★ **Zależność funkcyjna nietrywialna** – zależność funkcyjna w której co najmniej jeden z atrybutów typu B znajduje się pośród atrybutów typu A
  - Np.: tytuł, rok → rok, długość
- ★ **Nadklucz** – zbiór atrybutów który zawiera klucz.

# Przykład

## ★ Załóżmy, iż mamy zamodelować sytuację:

- każdy student może specjalizować się w kilku dziedzinach,
- student ma jednego asystenta w każdej dziedzinie,
- każda dziedzina ma kilku asystentów, ale jeden asystent doradza tylko w jednej dziedzinie,
- każdy asystent doradza kilku studentom w jednej dziedzinie.

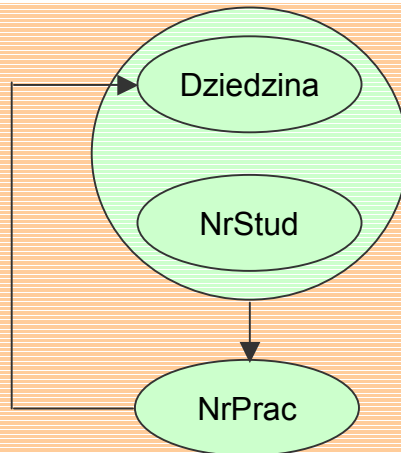


Specjalizacje (NrStud, Dzia dzina, NrPrac)

| Specjalizacje |                           |        |
|---------------|---------------------------|--------|
| NrStud        | Dzia dzina                | NrPrac |
| 123456        | Informatyka               | 234    |
| 234567        | Systemy informacyjne      | 345    |
| 345678        | Inżynieria oprogramowania | 456    |

## ★ Przedstawiony schemat spełnia wymogi 3PN

## Przykład c.d.



Specjalizacje (NrStud, Dziedzina, NrPrac)

| Specjalizacje |                           |        |
|---------------|---------------------------|--------|
| NrStud        | Dziedzina                 | NrPrac |
| 123456        | Informatyka               | 234    |
| 234567        | Systemy informacyjne      | 345    |
| 345678        | Inżynieria oprogramowania | 456    |

★ **Zauważmy że pomimo normalizacji do 3PN występują anomalie::**

- zmiana specjalizacji przez studenta 123456 powoduje utratę informacji o NrPrac 234,
- nie można wstawić informacji o NrPrac 789, który jest asystentem z Informatyki tak długo jak długo jakiś student nie wybierze tej specjalności,
- usunięcie studenta 345678 powoduje jednocześnie usunięcie informacji o Nr 456.

★ **Należy dokonać dekompozycji tabeli Specjalizacje:**

■ **Schemat 1**

AsystenciStudenta (NrStudenta, NrPrac)

DziedzinyAsystenta (NrPrac, Dziedzina)

■ **Schemat 2**

AsystenciStudenta (NrStudenta, Dziedzina)

DziedzinyAsystenta (NrPrac, Dziedzina)

## 6.11. Czwarta postać normalna (4PN)

**Relacja jest w czwartej postaci normalnej, wtedy i tylko wtedy gdy z nietrywialnych zależności wielowartościowych wynika, że pewien nadklucz wyznacza jakiś inny atrybut.**

- ★ **Relacja jest w 4PN, gdy zbiór atrybutów  $X$  określa wielowartościowo (niefunkcyjnie)  $Y$  to zachodzi jeden z następujących warunków:**
  - $Y$  jest puste lub zawiera się w  $X$ ,
  - Suma zbiorów  $X$  i  $Y$  jest kompletnym zbiorem atrybutów dla danej relacji,
  - $X$  zawiera klucz.
- ★ **Aby przejść z 3 PN do 4 PN, szukamy tabel, które zawierają dwie lub więcej niezależnych zależności wielowartościowych.**

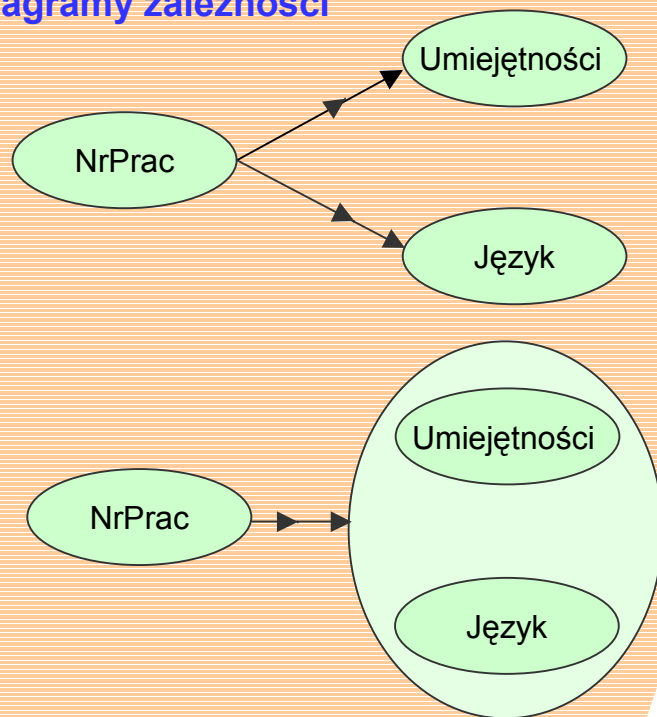
# Przykład

## ★ Załóżmy, iż mamy przechowywać:

- informację o pracownikach,
- umiejętnościach pracowników,
- znajomości języków przez pracowników.

- Pracownik może posiadać wiele umiejętności oraz znać kilka języków (umiejętności nie są związane ze znajomością języków).

### Diagramy zależności



### Tabele przed i po normalizacji do 4PN

| Pracownicy |                   |           |
|------------|-------------------|-----------|
| NrPrac     | Umiejętność       | Język     |
| 123456     | Obsługa komputera | angielski |
| 123456     | Obsługa komputera | francuski |
| 123456     | Prawo jazdy       | angielski |
| 234567     | Prawo jazdy       | niemiecki |
| 234567     | Obsługa komputera | angielski |

| Umiejętności |                   |
|--------------|-------------------|
| NrPrac       | Umiejętność       |
| 123456       | Obsługa komputera |
| 123456       | Prawo jazdy       |
| 234567       | Prawo jazdy       |
| 234567       | Obsługa komputera |

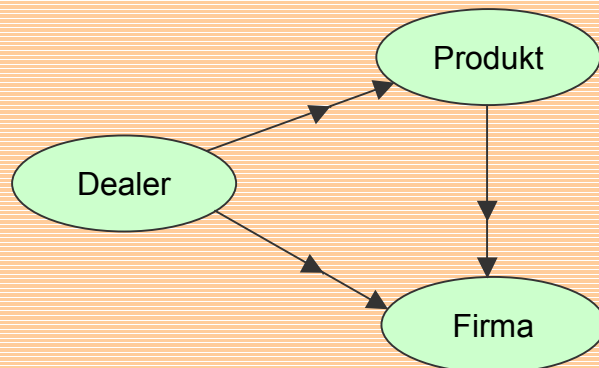
| Języki |           |
|--------|-----------|
| NrPrac | Język     |
| 123456 | angielski |
| 123456 | francuski |
| 234567 | niemiecki |
| 234567 | angielski |

## 6.12. Piątą postać normalna (5PN)

**Relacja jest w piątej postaci normalnej jeżeli jest w 4 PN i nie istnieje jej rozkład odwracalny na zbiór mniejszych tabel**

★ Załóżmy, iż mamy przechowywać:

- informację o dealerach samochodów,
- produktach sprzedawanych przez dealerów,



| PunktySprzedarzy |           |           |
|------------------|-----------|-----------|
| Dealer           | Producent | Typ       |
| Auto Zbyt        | Ford      | osobowy   |
| Auto Zbyt        | Opel      | dostawczy |
| Cztery Kółka     | Ford      | dostawczy |
| Cztery Kółka     | Opel      | osobowy   |

Dealerzy reprezentują firmy, firmy wytwarzają produkty i dealerzy sprzedają te produkty (diagram zależności). Nie można dokonać rozkładu struktury pokazanej na diagramie gdyż np.: Auto Zbyt sprzedaje samochody osobowe marki Ford i dostawcze marki Opel, a nie sprzedaje dostawczych fordów ani osobowych opeli.