# Bazy Danych i Usługi Sieciowe

## Model relacyjny

Paweł Daniluk

Wydział Fizyki

Jesień 2012



# Iloczyn kartezjański

## Iloczyn kartezjański zbiorów $A$, $B$

Zbiór wszystkich par uporządkowanych $(a, b)$, t. że $a \in A$ i $b \in B$.

$$(a, b) \in A \times B \iff a \in A \wedge b \in B$$

## Iloczyn kartezjański zbiorów $A_1, A_2, \ldots, A_n$

Zbiór wszystkich $n$-tek uporządkowanych $(a_1, a_2, \ldots, a_n)$, t. że $a_i \in A_i$.

$$(a_1, a_2, \ldots, a_n) \in A_1 \times A_2 \times \ldots \times A_n \iff \forall_{0<i\leq n} a_i \in A_i$$

# Relacje

- Relacja to podzbiór iloczynu kartezjańskiego
- Dwuargumentowa $R \subset A \times B$
- $n$-argumentowa $R \subset A_1 \times A_2 \times \ldots A_n$
- Reprezentacja jako dwuwymiarowa tabela

# Relacje

- Relacja to podzbiór iloczynu kartezjańskiego
- Dwuargumentowa $R \subset A \times B$
- $n$-argumentowa $R \subset A_1 \times A_2 \times \ldots A_n$
- Reprezentacja jako dwuwymiarowa tabela

## Relacja *Filmy*

| tytuł | rok | długość | typFilmu |
|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor |
| Potężne Kaczory | 1991 | 104 | kolor |
| Świat Wayne'a | 1992 | 95 | kolor |

# Relacje c.d.

## Atrybuty

*tytuł*, *rok*, *długość*, *typFilmu*

## Schemat

*Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*)

## Dziedziny

- *tytuł* – tekst
- *rok*, *długość* – liczby całkowite
- *typFilmu* – $\{\text{kolor}, \text{czarno-biały}\}$

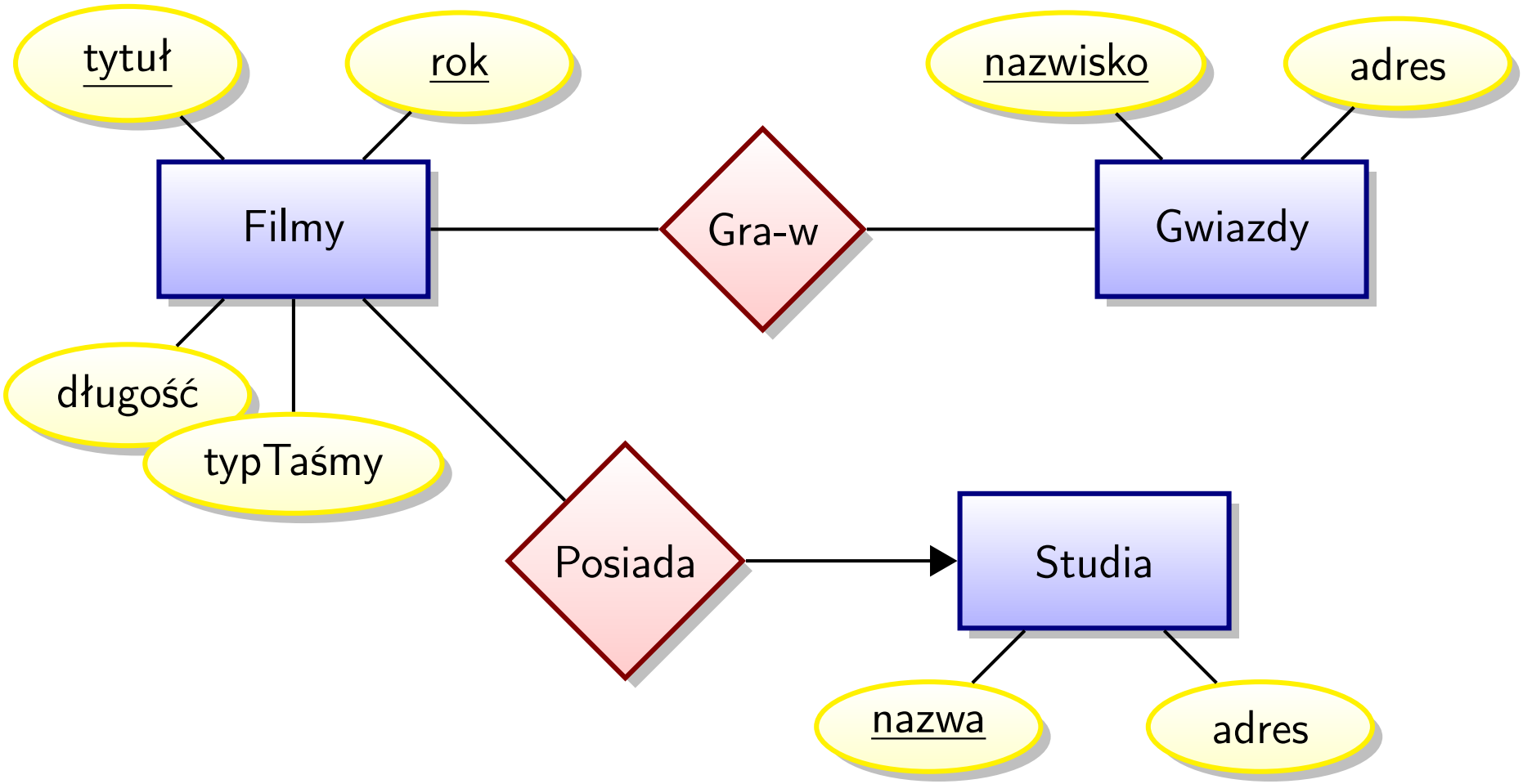
# Schemat

- Nazwy zbiorów w iloczynie kartezjańskim to *atrybuty*
- Nazwa relacji i zbiór uporządkowany jej atrybutów tworzą *schemat relacji*
- Projekt relacyjnej bazy danych zawiera zazwyczaj kilka relacji
- Zbiór schematów relacji projektu nazywamy *schematem bazy danych*
- Wiersze tabeli (poza nagłówkowym) – elementy relacji – nazywane są *krotkami*

# Dziedziny (typy danych)

- Każda składowa relacji ma swój typ danych
- Dostępne typy danych zależą od implementacji DBMS
- Najczęściej używane to
  - liczba całkowita
  - liczba rzeczywista
  - znak
  - łańcuch znaków
  - data
  - boolean

# Schemat związków encji



# Schemat relacyjny

## Zbiory encji $\rightarrow$ relacje

- *Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*)
- *Gwiazdy*(*nazwisko*, *adres*)
- *Studia*(*nazwa*, *adres*)

## Związki $\rightarrow$ relacje

- *Posiada*(*tytuł*, *rok*, *nazwaStudia*)
- *Gra-w*(*tytuł*, *rok*, *nazwiskoGwiazdy*)

# Relacje odpowiadające związkom

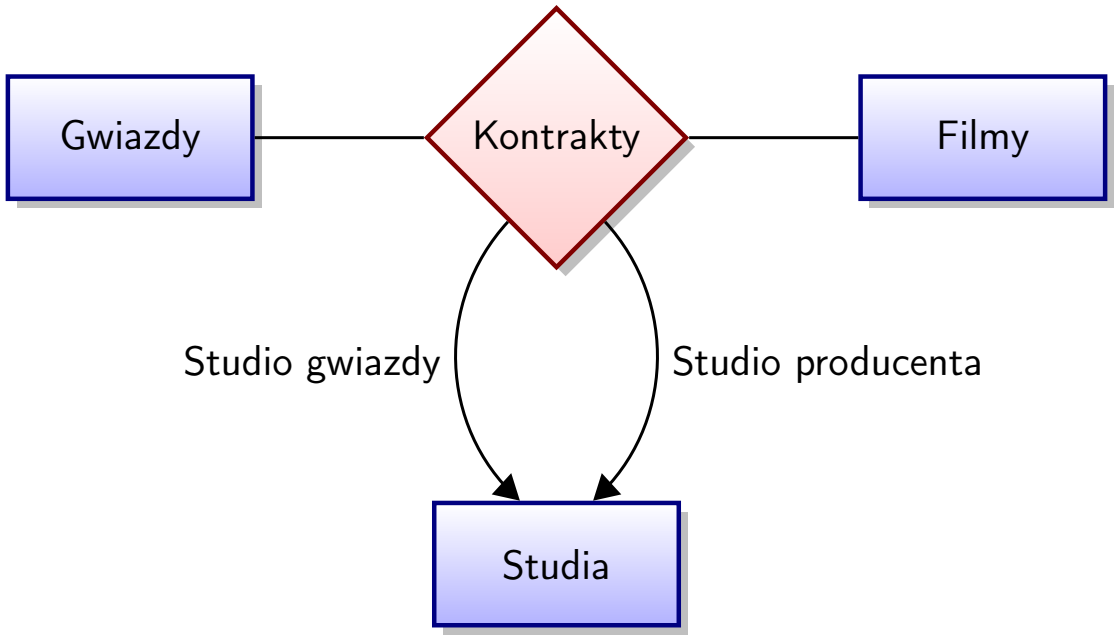
## *Posiada* – wiele do jeden

| tytuł | rok | nazwaStudia |
|---|---|---|
| Gwiezdne Wojny | 1977 | Fox |
| Potężne Kaczory | 1991 | Disney |
| Świat Wayne'a | 1992 | Paramount |

## *Gra-w* – wiele do wiele

| tytuł | rok | nazwiskoGwiazdy |
|---|---|---|
| Gwiezdne Wojny | 1977 | Carrie Fisher |
| Gwiezdne Wojny | 1977 | Mark Hamill |
| Gwiezdne Wojny | 1977 | Harrison Ford |
| Potężne Kaczory | 1991 | Emilio Estevez |
| Świat Wayne'a | 1992 | Dana Carvey |
| Świat Wayne'a | 1992 | Mike Meyers |

# Relacje odpowiadające związkom c.d.




*Kontrakty*(*nazwiskoGwiazdy*, *tytuł*, *rok*, *studioGwiazdy*, *studioProducenta*)

# Łączenie relacji

## *Filmy* i *Posiada*

*Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*, *nazwaStudia*)

| tytuł | rok | długość | typFilmu | nazwaStudia |
|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox |
| Potężne Kaczory | 1991 | 104 | kolor | Disney |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount |

# Łączenie relacji

## *Filmy* i *Posiada*

*Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*, *nazwaStudia*)

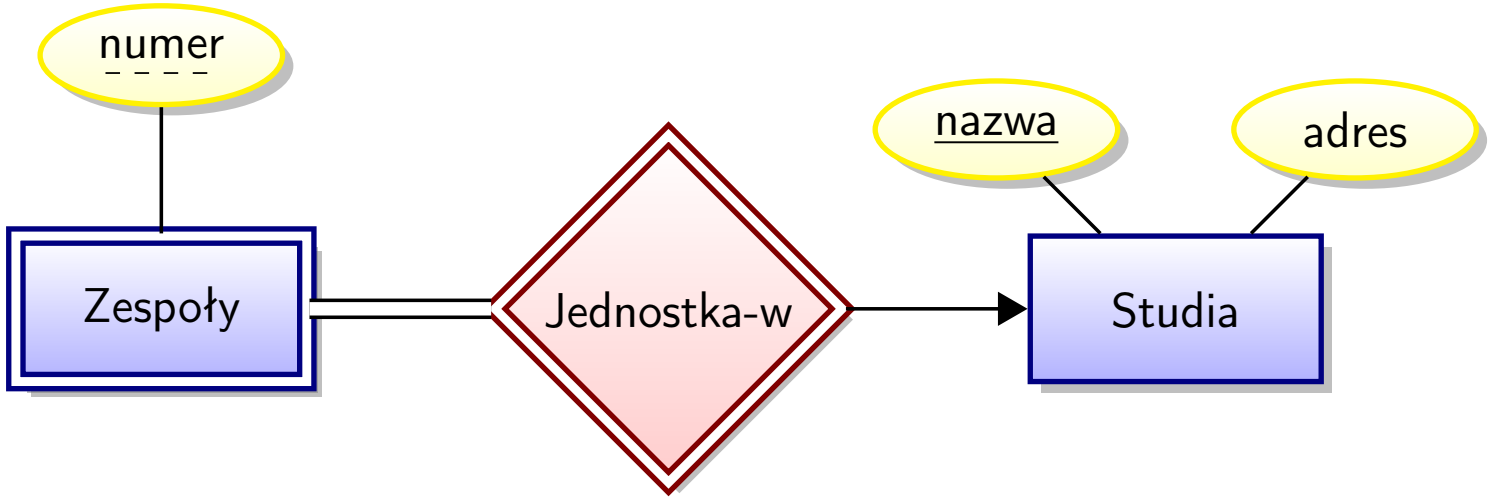
| tytuł | rok | długość | typFilmu | nazwaStudia |
|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox |
| Potężne Kaczory | 1991 | 104 | kolor | Disney |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount |

## Dołączenie *Gra-w* to zły pomysł

*Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*, *nazwaStudia*, *nazwiskoGwiazdy*)

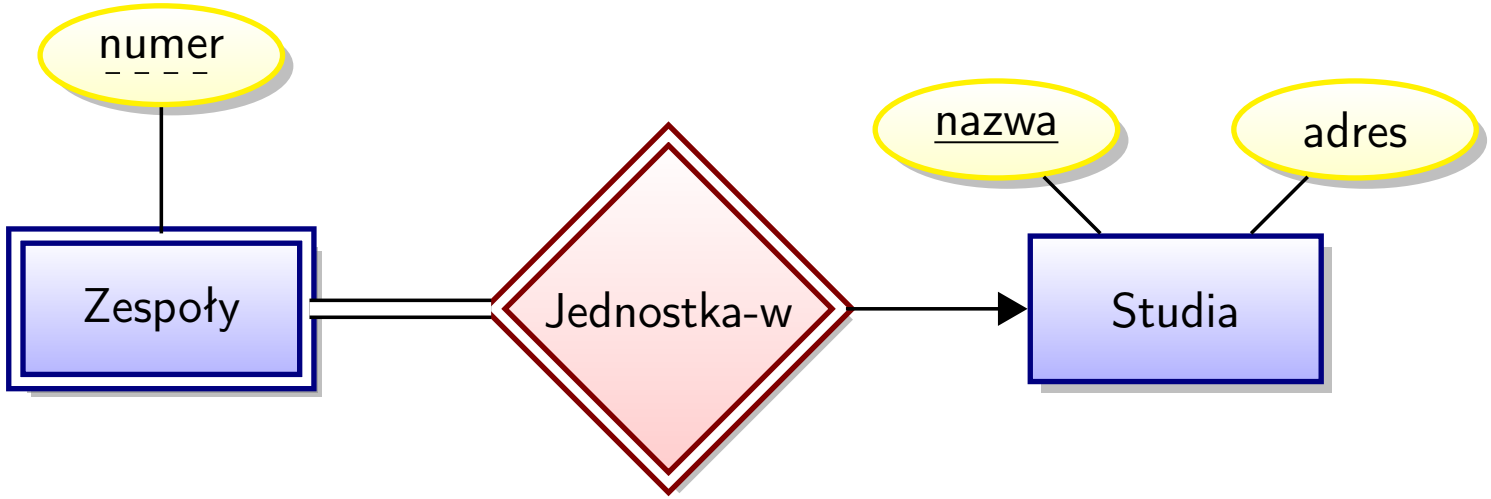
| tytuł | rok | długość | typFilmu | nazwaStudia | nazwiskoGwiazdy |
|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Carrie Fisher |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Mark Hamill |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Harrison Ford |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Emilio Estevez |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Dana Carvey |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Mike Meyers |

# Słabe encje



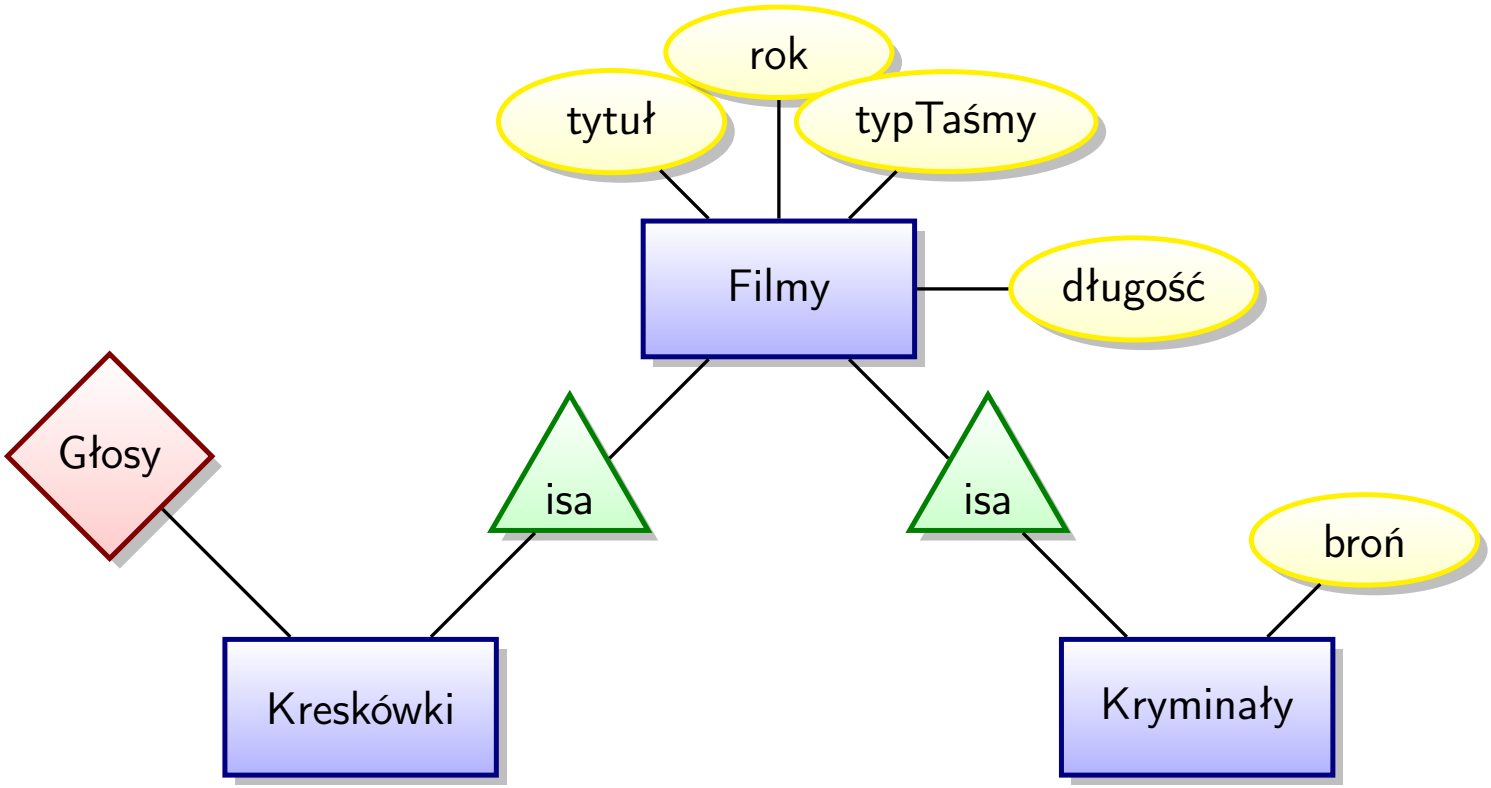

*Studia*(*nazwa*, *adres*, *rok*)
*Zespoły*(*numer*, *nazwaStudia*)
*Jednostka-w*(*numer*, *nazwaStudia*, *nazwa*)

# Słabe encje




*Studia*(*nazwa*, *adres*, *rok*)
*Zespoły*(*numer*, *nazwaStudia*)
*Jednostka-w*(*numer*, *nazwaStudia*, *nazwa*) nazwaStudia = nazwa

# Podklasy


rok
tytuł
typTaśmy
Filmy
długość
Głosy
isa
isa
broń
Kreskówki
Kryminały

# Podklasy c.d.

## Podejście E/R

- *Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*)
- *Kryminały*(*tytuł*, *rok*, *broń*)
- *Kreskówki*(*tytuł*, *rok*)

# Podklasy c.d.

## Podejście E/R

- *Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*)
- *Kryminały*(*tytuł*, *rok*, *broń*)
- *Kreskówki*(*tytuł*, *rok*)

## Podejście Obiektowe

- *Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*)
- *FilmyKres*(*tytuł*, *rok*, *długość*, *typFilmu*)
- *FilmyKrym*(*tytuł*, *rok*, *długość*, *typFilmu*, *broń*)
- *FilmyKrymKres*(*tytuł*, *rok*, *długość*, *typFilmu*, *broń*)

# Podklasy c.d.

## Podejście E/R

- *Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*)
- *Kryminały*(*tytuł*, *rok*, *broń*)
- *Kreskówki*(*tytuł*, *rok*)

## Podejście Obiektowe

- *Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*)
- *FilmyKres*(*tytuł*, *rok*, *długość*, *typFilmu*)
- *FilmyKrym*(*tytuł*, *rok*, *długość*, *typFilmu*, *broń*)
- *FilmyKrymKres*(*tytuł*, *rok*, *długość*, *typFilmu*, *broń*)

## Podejście z wartością NULL

- *Filmy*(*tytuł*, *rok*, *długość*, *typFilmu*, *broń*)

Atrybut *broń* może przyjmować wartość NULL. Relacja *Głosy* wyznacza bycie kreskówką.

# Zależności funkcyjne

## Zależność funkcyjna

*Jeśli dwie krotki relacji $R$ są zgodne dla atrybutów $A_1, A_2, \ldots, A_n$, to muszą być zgodne w pewnym innym atrybucie $B$*

## Notacja

$$A_1 A_2 \ldots A_n \to B$$

Wiele zależności funkcyjnych

$$A_1 A_2 \ldots A_n \to B_1$$
$$\ldots$$
$$A_1 A_2 \ldots A_n \to B_m$$

zapisujemy skrótowo jako

$$A_1 A_2 \ldots A_n \to B_1 \ldots B_m$$

# Zależności funkcyjne c.d.

| tytuł | rok | długość | typFilmu | nazwaStudia | nazwiskoGwiazdy |
|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Carrie Fisher |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Mark Hamill |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Harrison Ford |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Emilio Estevez |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Dana Carvey |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Mike Meyers |

*tytuł rok* $\rightarrow$ *długość typFilmu nazwaStudia*

Ta zależność nie zachodzi

*tytuł rok* $\rightarrow$ *nazwiskoGwiazdy*

# Klucze

- Atrybut lub zbiór atrybutów $\{A_1, A_2, \ldots, A_n\}$ tworzy *klucz relacji* $R$ jeśli
  1. Wszystkie pozostałe atrybuty relacji są funkcyjnie zależne od tych atrybutów,
  2. Nie istnieje podzbiór właściwy zbioru $\{A_1, A_2, \ldots, A_n\}$, od którego pozostałe atrybuty relacji $R$ są zależne funkcyjnie.
- Jeżeli klucz składa się z jednego atrybutu $A$, to mówimy, że $A$ jest kluczem.
- Dowolny zbiór atrybutów, którego podzbiorem jest klucz nazywany jest *nadkluczem*.

# Klucze c.d.

- Schemat relacji może zawierać kilka kluczy.
- W DBMS wyróżnia się jeden klucz jako podstawowy (*PRIMARY KEY*).
- Atrybuty należące do jakiegokolwiek klucza nazywa się *kluczowymi* lub *podstawowymi*.
- Atrybuty nie należące do żadnego klucza nazywa się *niekluczowymi* lub *wtórnymi*.

# Klucze c.d.

| tytuł | rok | długość | typFilmu | nazwaStudia | nazwiskoGwiazdy |
|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Carrie Fisher |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Mark Hamill |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Harrison Ford |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Emilio Estevez |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Dana Carvey |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Mike Meyers |

# Klucze c.d.

| tytuł | rok | długość | typFilmu | nazwaStudia | nazwiskoGwiazdy |
|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Carrie Fisher |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Mark Hamill |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Harrison Ford |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Emilio Estevez |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Dana Carvey |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Mike Meyers |

## Klucz?

*tytuł rok*

# Klucze c.d.

| tytuł | rok | długość | typFilmu | nazwaStudia | nazwiskoGwiazdy |
|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Carrie Fisher |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Mark Hamill |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Harrison Ford |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Emilio Estevez |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Dana Carvey |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Mike Meyers |

## Klucz

*tytuł rok nazwiskoGwiazdy*

# Wnioskowanie z zależności funkcyjnych

## Przykład – przechodniość

Jeżeli istnieją zależności funkcyjne $A \to B$ i $B \to C$, to istnieje zależność funkcyjna $A \to C$.

## Dowód

Weźmy dwie krotki: $t = (a, b_1, c_1)$, $u = (a, b_2, c_2)$

$$A \to B \implies b_1 = b_2$$

$$B \to C \implies c_1 = c_2$$

Każde dwie krotki zgodne dla $A$ są zgodne dla $C$.

$$A \to C$$

# Równoważność i wynikanie

- Dwa zbiory zależności funkcyjnych $S$ i $T$ są *równoważne*, jeśli zbiór instancji relacji spełniających $S$ jest równy zbiorowi instancji spełniających $T$.
- Zbiór zależności funkcyjnych $S$ *wynika* ze zbioru $T$, jeśli każda instancja spełniająca wszystkie zależności funkcyjne z $T$ spełnia także zależności z $S$.

**Przypomnienie**

Zbiór zależności funkcyjnych stanowi część schematu bazy danych.

# Reguły podziału i łączenia

## Reguła podziału

Zależność funkcyjna $A_1A_2 \ldots A_n \to B_1B_2 \ldots B_m$ można zastąpić zbiorem zależności funkcyjnych $A_1A_2 \ldots A_n \to B_i$, $i = 1, 2, \ldots, m$.

## Reguła łączenia

Zbiór zależności funkcyjnych $A_1A_2 \ldots A_n \to B_i$, $i = 1, 2, \ldots, m$ można zastąpić pojedynczą zależnością funkcyjną $A_1A_2 \ldots A_n \to B_1B_2 \ldots B_m$.

## Przykład

$$\begin{array}{l} \textit{tytuł rok} \to \textit{długość typFilmu} \\ \textit{tytuł rok} \to \textit{długość typFilmu} \\ \textit{tytuł rok} \to \textit{nazwaStudia} \end{array} \quad \Longrightarrow \quad \begin{array}{l} \textit{tytuł rok} \to \textit{długość typFilmu nazwaStudia} \\ \textit{tytuł rok} \to \textit{długość typFilmu} \\ \ldots \end{array}$$

# Reguły podziału i łączenia

## Reguła podziału

Zależność funkcyjna $A_1A_2 \ldots A_n \to B_1B_2 \ldots B_m$ można zastąpić zbiorem zależności funkcyjnych $A_1A_2 \ldots A_n \to B_i$, $i = 1, 2, \ldots, m$.

## Reguła łączenia

Zbiór zależności funkcyjnych $A_1A_2 \ldots A_n \to B_i$, $i = 1, 2, \ldots, m$ można zastąpić pojedynczą zależnością funkcyjną $A_1A_2 \ldots A_n \to B_1B_2 \ldots B_m$.

## Przykład

$$\begin{array}{l} \textit{tytuł rok} \to \textit{długość typFilmu} \\ \textit{tytuł rok} \to \textit{długość typFilmu} \\ \textit{tytuł rok} \to \textit{nazwaStudia} \end{array} \quad \Longrightarrow \quad \begin{array}{l} \textit{tytuł rok} \to \textit{długość typFilmu nazwaStudia} \\ \textit{tytuł rok} \to \textit{długość typFilmu} \\ \ldots \end{array}$$

## Ta implikacja nie zachodzi

$$\textit{tytuł rok} \to \textit{długość} \Longrightarrow \begin{array}{l} \textit{tytuł} \to \textit{długość} \\ \textit{rok} \to \textit{długość} \end{array}$$

# Trywialne zależności funkcyjne

Zależność funkcyjna $A_1 A_2 \ldots A_n \to B_1 B_2 \ldots B_m$ jest

- *trywialna*, jeśli zbiór atrybutów typu $B$ jest podzbiorem zbioru atrybutów typu $A$;
- *nietrywialna*, jeśli co najmniej jeden z atrybutów typu $B$ nie znajduje się sie pośród atrybutów typu $A$;
- *całkowicie nietrywialna*, jeśli żaden z atrybutów typu $B$ nie znajduje sie pośród atrybutów typu $A$.

## Reguła zależności trywialnych

Zależność funkcyjna $A_1 A_2 \ldots A_n \to B_1 B_2 \ldots B_m$ jest równoważna zależności $A_1 A_2 \ldots A_n \to C_1 C_2 \ldots C_k$, gdzie $C$ są tymi elementami z $B$, które nie należą do $A$.

# Trywialne zależności funkcyjne c.d.

## Przykład

*tytuł rok* $\rightarrow$ *tytuł rok* – trywialna
*tytuł rok* $\rightarrow$ *długość rok* – nietrywialna
*tytuł rok* $\rightarrow$ *długość typFilmu* – całkowicie nietrywialna

# Domknięcie zbioru atrybutów

*Domknięciem* zbioru atrybutów $\{A_1, A_2, \ldots, A_n\}$ nad zbiorem zależności funkcyjnych $S$ nazywamy taki zbiór atrybutów $B$, że jeżeli pewna relacja $R$ spełnia wszystkie zależności ze zbioru $S$, to spełnia także zależność $A_1 A_2 \ldots A_n \rightarrow B$.

## Notacja

$$\{A_1, A_2, \ldots, A_n\}^+$$

# Algorytm wyznaczania domknięcia

1. $X := \{A_1, A_2, \ldots, A_n\}$
2. Znajdujemy wszystkie zależności funkcyjne postaci $B_1 B_2 \ldots B_m \to C$, dla których $B_1, B_2, \ldots, B_m$ należą do zbioru atrubutów $X$, a $C$ nie należy. Wówczas dołączmy $C$ do $X$

$$X := X \cup \{C\}$$

3. Powtarzamy krok 2 tak długo, jak długo nie będzie można dołączyć do $X$ żadnego nowego atrybutu.

# Algorytm wyznaczania domknięcia

## Przykład

Relacja: $R(A, B, C, D, E, F)$
Zależności funkcyjne: $S = \{AB \to C, BC \to AD, D \to E, CF \to B\}$
Obliczymy $\{A, B\}^+$:

1. $X := \{A, B\}$

# Algorytm wyznaczania domknięcia

## Przykład

Relacja: $R(A, B, C, D, E, F)$
Zależności funkcyjne: $S = \{AB \to C, BC \to AD, D \to E, CF \to B\}$
Obliczymy $\{A, B\}^+$:

1. $X := \{A, B\}$
2. $A, B \in X$, $AB \to C$, więc $X := X \cup \{C\} = \{A, B, C\}$

# Algorytm wyznaczania domknięcia

## Przykład

Relacja: $R(A, B, C, D, E, F)$
Zależności funkcyjne: $S = \{AB \to C, BC \to AD, D \to E, CF \to B\}$
Obliczymy $\{A, B\}^+$:

1. $X := \{A, B\}$
2. $A, B \in X$, $AB \to C$, więc $X := X \cup \{C\} = \{A, B, C\}$
3. $B, C \in X$, $BC \to AD$, więc $X := X \cup \{D\} = \{A, B, C, D\}$

# Algorytm wyznaczania domknięcia

## Przykład

Relacja: $R(A, B, C, D, E, F)$
Zależności funkcyjne: $S = \{AB \to C, BC \to AD, D \to E, CF \to B\}$
Obliczymy $\{A, B\}^+$:

1. $X := \{A, B\}$
2. $A, B \in X$, $AB \to C$, więc $X := X \cup \{C\} = \{A, B, C\}$
3. $B, C \in X$, $BC \to AD$, więc $X := X \cup \{D\} = \{A, B, C, D\}$
4. $D \in X$, $D \to E$, więc $X := X \cup \{E\} = \{A, B, C, D, E\}$

# Algorytm wyznaczania domknięcia

## Przykład

Relacja: $R(A, B, C, D, E, F)$
Zależności funkcyjne: $S = \{AB \to C, BC \to AD, D \to E, CF \to B\}$
Obliczymy $\{A, B\}^+$:

1. $X := \{A, B\}$
2. $A, B \in X$, $AB \to C$, więc $X := X \cup \{C\} = \{A, B, C\}$
3. $B, C \in X$, $BC \to AD$, więc $X := X \cup \{D\} = \{A, B, C, D\}$
4. $D \in X$, $D \to E$, więc $X := X \cup \{E\} = \{A, B, C, D, E\}$

# Algorytm wyznaczania domknięcia

## Przykład

Relacja: $R(A, B, C, D, E, F)$
Zależności funkcyjne: $S = \{AB \to C, BC \to AD, D \to E, CF \to B\}$
Obliczymy $\{A, B\}^+$:

1. $X := \{A, B\}$
2. $A, B \in X$, $AB \to C$, więc $X := X \cup \{C\} = \{A, B, C\}$
3. $B, C \in X$, $BC \to AD$, więc $X := X \cup \{D\} = \{A, B, C, D\}$
4. $D \in X$, $D \to E$, więc $X := X \cup \{E\} = \{A, B, C, D, E\}$

Lewa strona $CF \to B$ nie znajdzie się w $X$.

# Algorytm wyznaczania domknięcia

## Przykład

Relacja: $R(A, B, C, D, E, F)$
Zależności funkcyjne: $S = \{AB \to C, BC \to AD, D \to E, CF \to B\}$
Obliczymy $\{A, B\}^+$:

1. $X := \{A, B\}$
2. $A, B \in X$, $AB \to C$, więc $X := X \cup \{C\} = \{A, B, C\}$
3. $B, C \in X$, $BC \to AD$, więc $X := X \cup \{D\} = \{A, B, C, D\}$
4. $D \in X$, $D \to E$, więc $X := X \cup \{E\} = \{A, B, C, D, E\}$

Lewa strona $CF \to B$ nie znajdzie się w $X$.

$$\{A, B\}^+ = \{A, B, C, D, E\}$$

# Rzutowanie zależności funkcyjnych

## Dane

Relacja: $R(A, B, C, D)$
Zależności funkcyjne: $S = \{A \to B, B \to C, C \to D\}$

## Zadanie

Wyeliminować atrybut $B$.

## Wynik

Relacja: $R(A, C, D)$
Zależności funkcyjne: $S = \{A \to C, C \to D\}$

# Anomalie

## *Filmy*

| tytuł | rok | długość | typFilmu | nazwaStudia | nazwiskoGwiazdy |
|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Carrie Fisher |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Mark Hamill |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Harrison Ford |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Emilio Estevez |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Dana Carvey |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Mike Meyers |

- Redundancja
- Anomalie przy aktualizacji
- Anomalie przy usuwaniu
- Anomalie przy wstawianiu

# Normalizacja

- Wyeliminowanie redundancji
- Wyeliminowanie anomalii
- *Dekompozycja* - rozkład relacji na mniejsze, które nie mają redundancji i anomalii
- Potrzebna informacja o zależnościach funkcyjnych

# Własności dla normalizacji

- Zachowanie atrybutów
- Zachowanie informacji
- Zachowanie zależności funkcyjnych w relacjach po normalizacji

# Pierwsza postać normalna

## 1NF

Schemat relacji $R$ znajduje się w pierwszej postaci normalnej (1NF), jeżeli wartości atrybutów są atomowe (niepodzielne).

## Przykład

| tytuł | rok | długość | typFilmu | nazwaStudia | gwiazdy |
|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Carrie Fisher, Mark Hamill, Harrison Ford |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Emilio Estevez |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Dana Carvey, Mike Meyers |

Filmy(tytuł, rok, długość, nazwaStudia, Gwiazdy(nazwisko))

# Dekompozycja do 1NF

## Algorytm

- Załóżmy, że relacja $R$ zawiera relację zagnieżdżoną $P$
- Tworzymy relację $R'$ dla relacji zewnętrznej $R$ z atrybutów nie zawierających atrybutu złożonego zawierającego relację $P$.
- Tworzymy relację $P'$ dla relacji wewnętrznej, do której dodajemy klucz relacji zewnętrznej.
- Kluczem nowej relacji $P'$ jest suma klucza relacji zewnętrznej $R$ i klucza relacji wewnętrznej $P'$.

## Przykład

$R$: Filmy(tytuł, rok, długość, nazwaStudia, Gwiazdy(nazwisko))
$P$: Gwiazdy(nazwisko)

$R'$: Filmy(tytuł, rok, długość, nazwaStudia)
$P'$: Gwiazdy(tytuł, rok, nazwisko)

# Dekompozycja do 1NF c.d.

## *Filmy* – przed normalizacją

| tytuł | rok | długość | typFilmu | nazwaStudia | gwiazdy |
|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Carrie Fisher, Mark Hamill, Harrison Ford |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Emilio Estevez |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Dana Carvey, Mike Meyers |

## Po normalizacji

### *Filmy*

| tytuł | rok | długość | typFilmu | nazwaStudia |
|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox |
| Potężne Kaczory | 1991 | 104 | kolor | Disney |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount |

### *Gwiazdy*

| tytuł | rok | nazwiskoGwiazdy |
|---|---|---|
| Gwiezdne Wojny | 1977 | Carrie Fisher |
| Gwiezdne Wojny | 1977 | Mark Hamill |
| Gwiezdne Wojny | 1977 | Harrison Ford |
| Potężne Kaczory | 1991 | Emilio Estevez |
| Świat Wayne'a | 1992 | Dana Carvey |
| Świat Wayne'a | 1992 | Mike Meyers |

# Częściowe zależności funkcyjne

- Zbiór atrybutów $B$ jest *w pełni funkcyjnie zależny* od zbioru atrybutów $A$, jeżeli $A \to B$ i nie istnieje podzbiór właściwy $A' \subset A$ taki, że $A' \to B$.
- Zbiór atrybutów $B$ jest *częściowo funkcyjnie zależny* od zbioru atrybutów $A$, jeżeli $A \to B$ i istnieje podzbiór właściwy $A' \subset A$ taki, że $A' \to B$.

# Druga postać normalna - 2NF

## 2NF

Schemat relacji $R$ znajduje się w drugiej postaci normalnej (2NF), jeżeli żaden atrybut niekluczowy tej relacji nie jest częściowo zależny funkcyjnie od żadnego z kluczy relacji $R$.

## *Filmy*

| tytuł | rok | długość | typFilmu | nazwaStudia | nazwiskoGwiazdy | adresGwiazdy |
|---|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Carrie Fisher | Malibu |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Mark Hamill | Brentwood |
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Harrison Ford | Hollywood |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Emilio Estevez | West Hollywood |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Dana Carvey | Beverly Hills |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Mike Meyers | Carlsbad |

## Zależności funkcyjne

*tytuł rok nazwiskoGwiazdy* $\rightarrow$ *długość typFilmu nazwaStudia adresGwiazdy*
*tytuł rok* $\rightarrow$ *długość typFilmu nazwaStudia*
*nazwiskoGwiazdy* $\rightarrow$ *adresGwiazdy*

# Zależności przechodnie

- Zbiór atrybutów $B$ jest *przechodnio funkcyjnie zależny* od zbioru atrybutów $A$, jeżeli $A \to B$ i istnieje zbiór atrybutów $C$, nie będący podzbiorem żadnego klucza, taki, że $A \to C$ i $C \to B$.
- Zależność fukncyjna $A \to B$ jest *zależnością przechodnią* jeżeli istnieje podzbiór atrybutów $C$ taki, że zachodzi $A \to C$, $C \to B$ i nie zachodzi $C \to A$ lub $B \to C$.

# Trzecia postać normalna - 3NF

## 3NF

Schemat relacji $R$ znajduje się w trzeciej postaci normalnej (3NF), jeżeli dla każdej zależności funkcyjnej $\{A_1, \ldots, A_n\} \rightarrow B$ zbiór $\{A_1, \ldots, A_n\}$ jest nadkluczem lub $B$ jest elementem pewnego klucza.

## *StudiaFilmowe*

| tytuł | rok | długość | typFilmu | nazwaStudia | adresStudia |
|---|---|---|---|---|---|
| Gwiezdne Wojny | 1977 | 124 | kolor | Fox | Hollywood |
| Potężne Kaczory | 1991 | 104 | kolor | Disney | Buena Vista |
| Świat Wayne'a | 1992 | 95 | kolor | Paramount | Hollywood |

## Zależności funkcyjne

*tytuł rok* $\rightarrow$ *długość typFilmu nazwaStudia adresStudia*
*nazwaStudia* $\rightarrow$ *adresStudia*

# Postać normalna Boyce'a-Codda - BCNF

## BCNF

Schemat relacji $R$ znajduje się w postaci normalnej Boyce'a-Codda (BCNF), jeżeli dla każdej zależności funkcyjnej $\{A_1, \ldots, A_n\} \rightarrow B$ zbiór $\{A_1, \ldots, A_n\}$ jest nadkluczem.

- Warunek BCNF jest dostateczny dla 3NF, ale nie konieczny.
- BCNF $\Longrightarrow$ 3NF

# Różnica pomiędzy BCNF a 3NF

## *Projekcje*

| kino | miasto | tytuł |
|---|---|---|
| Polonia | Warszawa | Gwiezdne Wojny |
| Moskwa | Warszawa | Potężne Kaczory |
| Dom Kultury Kolejarza "Stokrotka" | Kogutkowo Górne | Gwiezdne Wojny |

## Zależności funkcyjne

$kino \rightarrow miasto$
$tytuł\ miasto \rightarrow kino$

## Klucze

(*tytuł*, *miasto*)
(*kino*, *tytuł*)

3NF, ale nie BCNF

## Dekompozycja do BCNF

Kina(*kino*, *miasto*)
Projekcje(*kino*, *tytuł*)
Zależność $tytuł\ miasto \rightarrow kino$ nie jest zachowana.