

## Wyjaśnienie zależności funkcyjnych oraz kluczy

### POJĘCIA

**Atrybut KLUCZOWY** (ang. prime attribute) to taki atrybut tabeli, który wchodzi w skład dowolnego klucza.

**Atrybut NIEKLUCZOWY** (ang. non-prime attribute) to taki atrybut tabeli, który nie wchodzi w skład żadnego klucza. Atrybut niekluczowy nazywany jest także atrybutem informacyjnym.

**KLUCZ** (ang. key) albo **KLUCZ KANDYDUJĄCY** (ang. candidate key) jest to **minimalny** podzbiór atrybutów relacji, który pozwala jednoznacznie identyfikować każdą krotkę.

**NADKLUCZ** jest to **dowolny** (niekoniecznie minimalny) podzbiór atrybutów relacji, który pozwala identyfikować jednoznacznie każdą krotkę.

**KLUCZ PODSTAWOWY** albo **KLUCZ GŁÓWNY** (ang. primary key) jest to jeden dowolny klucz wybrany spośród kluczy kandydujących.

Dla przykładowej relacji **PODATNICY(NIP, PESEL, NAZWISKO, IMIĘ)**:

- kluczem (kluczem kandydującym) jest NIP
- kluczem (kluczem kandydującym) jest PESEL
- nadkluczem jest (NIP, PESEL)
- nadkluczem jest (NIP, PESEL, NAZWISKO)
- nadkluczem jest (NIP, PESEL, IMIĘ)
- nadkluczem jest (NIP, PESEL, NAZWISKO, IMIĘ)
- atrybutem kluczowym jest NIP
- atrybutem kluczowym jest PESEL
- atrybutem niekluczowym jest IMIĘ
- atrybutem niekluczowym jest NAZWISKO

**Zależność funkcyjną** w postaci  $A \rightarrow B$  należy odczytywać (rozumieć):

- A determinuje B
- A określa B
- B zależy funkcyjnie od A

**Prawa (aksjomaty) Armstronga:**

(A1) Zwrotność: Jeżeli  $Y \subseteq X$  to  $X \rightarrow Y$

(A2) Rozszerzanie: Jeżeli  $X \rightarrow Y$  to zachodzi  $XZ \rightarrow YZ$

(A3) Przechodniość: Jeżeli  $X \rightarrow Y$  oraz  $Y \rightarrow Z$  to zachodzi  $X \rightarrow Z$

**Domknięcie (zbioru atrybutów relacji nad zbiorem zależności funkcyjnych)**

Niech dany będzie zbiór atrybutów  $A = (A_1, A_2, \dots, A_n)$  oraz zbiór zależności funkcyjnych  $F = (F_1, F_2, \dots, F_n)$ . Domknięciem zbioru atrybutów  $A$  nad zbiorem zależności  $F$  nazywamy taki zbiór atrybutów  $B$ , w którym dla każdego atrybutu  $B_i$ , należącego do pewnej relacji  $R$  spełniającej zależności  $F$ , spełniona jest zależność  $A_1 A_2 \dots A_n \rightarrow B_i$ .

## PRZYKŁADY

### Przykład 1

Dana jest relacja 3-atrybutowa  $R(A, B, C)$  dla której zachodzą następujące zależności funkcyjne  $F = \langle A \rightarrow B, B \rightarrow C \rangle$

Dla wymienionych wyżej zależności tworzymy tabelę, w której w kolumnie "Lewostronne" umieszczamy te atrybuty, które we wszystkich zależnościach funkcyjnych występują **tylko po lewej** stronie a także te, które w ogóle w zależnościach funkcyjnych nie występują tj. znajdują się w relacji  $R$  lecz brak ich w zależnościach funkcyjnych. W kolumnie "Obustronne" umieszczamy te atrybuty, które we zależnościach funkcyjnych występują **po obu** stronach (np. w jednej zależności po stronie prawej a w innej po stronie lewej). W kolumnie "Prawostronne" umieszczamy te atrybuty, które we wszystkich zależnościach funkcyjnych występują **tylko po prawej** stronie.

Dla zależności funkcyjnych z przykładu tj.  $F = \langle A \rightarrow B, B \rightarrow C \rangle$  otrzymujemy:

Lewostronne atrybuty kluczowe	Obustronne atrybuty potencjalnie kluczowe	Prawostronne atrybuty niekluczowe
A	B	C

Stąd wnioskujemy, że atrybuty które znalazły się w kolumnie "Lewostronne" **SĄ** na pewno atrybutami kluczowymi (muszą się pojawić w **każdym** kluczu). Atrybuty, które znalazły się w kolumnie "Prawostronne" **NIE SĄ** na pewno atrybutami kluczowymi (nie mogą się pojawić w **żadnym** kluczu). Natomiast atrybuty, które znajdują się w kolumnie "Obustronne" mogą ale nie muszą być atrybutami kluczowymi (mogą się pojawić w jednym kluczu ALBO mogą się pojawić w kilku kluczach ALBO mogą się pojawić we wszystkich kluczach ALBO mogą się nie pojawić w żadnym kluczu).

Jak sprawdzić czy atrybuty znajdujące się w kolumnie "Obustronne" są lub nie są atrybutami kluczowymi? Jeżeli są, to w jakich kluczach?

W tym celu musimy wyznaczyć domknięcie dla zbioru atrybutów kluczowych. W naszym przypadku dla atrybutu  $A$ , który jest na pewno atrybutem wchodzącym w skład klucza, musimy wyznaczyć **domknięcie**, czyli  $A^+$ . W tym celu będziemy korzystać z praw Armstronga:

$A^+ \rightarrow ?$

$A^+ \rightarrow A$  (zgodnie z prawem zwrotności Armstronga, gdyż  $A \subseteq A^+$ )

$A^+ \rightarrow AB$  (gdyż  $A$  determinuje  $B$  tj. istnieje zależność funkcyjna  $A \rightarrow B$ )

$A^+ \rightarrow ABC$  (gdyż  $B$  determinuje  $C$  tj. istnieje zależność funkcyjna  $B \rightarrow C$ )

Ponieważ po prawej stronie pojawiły się wszystkie atrybuty relacji  $R(A, B, C)$  to możemy powiedzieć, że atrybut  $A$  determinuje (określa) **wszystkie** atrybuty  $(A, B, C)$  relacji  $R$ . Jest to warunek konieczny, aby uznać, że określony atrybut jest **kluczem** relacji.

Komentarz: Powyżej powiedzieliśmy, że gdy pojedynczy atrybut lub zbiór atrybutów determinuje wszystkie atrybuty relacji to możemy uznać go za klucz tej relacji. W znakomitej większości przypadków tak będzie, jednak w bardzo szczególnych przypadkach może się okazać, że otrzymamy nie klucz lecz nadklucz! Dlatego w takich przypadkach, dla pewności powinniśmy sprawdzić, czy otrzymany atrybut albo zbiór atrybutów spełnia warunek klucza tj. czy jest zbiorem minimalnym? Dla wyżej przedstawionego przykładu możemy powiedzieć, że ponieważ atrybut  $A$  jest atrybutem pojedynczym, w związku z tym możemy na pewno stwierdzić, że jest minimalnym podzbiorem atrybutów relacji  $R$  (nie ma mniejszego podzbioru atrybutów relacji  $R$  niż pozbiór 1-elementowy). Zatem  $A$  na pewno jest kluczem.

Czy istnieją inne klucze w zadanej relacji? Formalnie powinniśmy sprawdzić domknięcie dla pozostałych, potencjalnych atrybutów, które znajdują się w kolumnie "Obustronne" tj. powinniśmy sprawdzić  $AB^+ \rightarrow ?$  oraz  $AC^+ \rightarrow ?$  Ale czy takie sprawdzenie na pewno jest konieczne? Przecież z definicji klucza wynika, że

klucz to minimalny zbiór atrybutów. Zatem biorąc pod uwagę wyznaczony wcześniej klucz A możemy stwierdzić, że:

- AB na pewno nie jest kluczem w relacji R, gdyż mniejszym podzbiorem ze zbioru AB, który jest już kluczem, jest pojedynczy atrybut A (AB jest zbiorem 2-elementowym, natomiast A jest zbiorem 1-elementowym);
- AC na pewno nie jest kluczem w relacji R, gdyż mniejszym podzbiorem ze zbioru AB, który jest już kluczem, jest pojedynczy atrybut A (AC jest zbiorem 2-elementowym, natomiast A jest zbiorem 1-elementowym);

AB oraz AC są nadkluczami w relacji R.

Na marginesie, dlaczego A pojawia się w obu domknięciach? Pojawia się dlatego, gdyż A jest **na pewno** atrybutem kluczowym, zatem musi się pojawić w **każdym** kluczu, natomiast B i C mogą się pojawić ale nie muszą.

Wynik do sprawdzenia tutaj: [http://uisacad5.uis.edu/cgi-bin/mcrem2/database\\_design\\_tool.cgi](http://uisacad5.uis.edu/cgi-bin/mcrem2/database_design_tool.cgi)

## Przykład 2

Dana jest relacja 4-atrybutowa  $R(K, L, M, N)$  dla której zachodzą następujące zależności funkcyjne  $F = \langle MN \rightarrow K, K \rightarrow L, L \rightarrow M \rangle$

Dla zależności funkcyjnych z niniejszego przykładu tj.  $F = \langle \text{MN} \rightarrow \text{K}, \text{K} \rightarrow \text{L}, \text{L} \rightarrow \text{M} \rangle$  otrzymujemy:

Lewostronne atrybuty kluczowe	Obustronne atrybuty potencjalnie kluczowe	Prawostronne atrybuty niekluczowe
N	MKL	-

Atrybutem kluczowym jest na pewno N, natomiast M, K oraz L są potencjalnymi atrybutami kluczowymi.

Wyznaczamy domknięcie dla zbioru atrybutów kluczowych tj. dla N.

$N^+ \rightarrow ?$

$N^+ \rightarrow \text{N}$  (zgodnie z prawem zwrotności Armstronga, gdyż  $N \subseteq N^+$ )

Tutaj się zatrzymujemy, gdyż stosując prawa Armstronga nic więcej zrobić nie możemy. Możemy zatem wysnuć wniosek, że N nie jest kluczem (po prawej stronie nie otrzymaliśmy listy wszystkich atrybutów relacji R).

Zatem zamiast rozpatrywać tylko N, dołożmy do tego atrybutu jeden z atrybutów, które potencjalnie mogą być atrybutami kluczowymi tj. M, K, L. Będziemy zatem musieli rozważyć 3 przypadki: MN, KN, LN.

Przypadek 1)

Zatem wyznaczmy domknięcie dla zbioru atrybutów kluczowych w postaci MN.

$MN^+ \rightarrow ?$

$MN^+ \rightarrow \text{MN}$  (zgodnie z prawem zwrotności Armstronga, gdyż  $MN \subseteq MN^+$ )

$MN^+ \rightarrow \text{MNK}$  (gdyż MN determinuje K tj. istnieje zależność funkcyjna  $MN \rightarrow K$ )

$MN^+ \rightarrow \text{MNKL}$  (gdyż K determinuje L tj. istnieje zależność funkcyjna  $K \rightarrow L$ )

Komentarz:  $\text{NM} \rightarrow \text{K}$  jest tożsame dla  $\text{MN} \rightarrow \text{K}$  (kolejność atrybutów nie ma znaczenia)

Ponieważ po prawej stronie pojawiły się wszystkie atrybuty relacji  $R(K, L, M, N)$  to możemy powiedzieć, że atrybuty MN determinują (określa) **wszystkie** atrybuty (K, L, M, N) relacji R. Jest to warunek konieczny, aby uznać atrybuty MN jako **klucz** relacji.

Następnie wyznaczmy domknięcie dla zbioru atrybutów kluczowych w postaci KN.

$KN^+ \rightarrow ?$

$KN^+ \rightarrow \text{KN}$  (zgodnie z prawem zwrotności Armstronga, gdyż  $KN^+ \subseteq KN$ )

$KN^+ \rightarrow \text{KNL}$  (gdyż K determinuje L tj. istnieje zależność funkcyjna  $K \rightarrow L$ )

$KN^+ \rightarrow \text{KNLM}$  (gdyż L determinuje M tj. istnieje zależność funkcyjna  $L \rightarrow M$ )

Ponieważ po prawej stronie pojawiły się wszystkie atrybuty relacji  $R(K, L, M, N)$  to możemy powiedzieć, że atrybuty KN determinują (określa) **wszystkie** atrybuty (K, L, M, N) relacji R. Jest to warunek konieczny, aby uznać atrybuty KN jako **klucz** relacji.

Przypadek 2)

Następnie wyznaczmy domknięcie dla zbioru atrybutów kluczowych w postaci KN.

$KN^+ \rightarrow ?$

$KN^+ \rightarrow \text{KN}$  (zgodnie z prawem zwrotności Armstronga, gdyż  $KN^+ \subseteq KN$ )

$KN^+ \rightarrow \text{KNL}$  (gdyż K determinuje L tj. istnieje zależność funkcyjna  $K \rightarrow L$ )

$KN^+ \rightarrow \text{KNLM}$  (gdyż L determinuje M tj. istnieje zależność funkcyjna  $L \rightarrow M$ )

Ponieważ po prawej stronie pojawiły się wszystkie atrybuty relacji  $R(K, L, M, N)$  to możemy powiedzieć, że atrybuty KN determinują (określa) **wszystkie** atrybuty (K, L, M, N) relacji R. Jest to warunek konieczny, aby uznać atrybuty KN jako **klucz** relacji.

Przypadek 3)

Na koniec wyznaczmy domknięcie dla zbioru atrybutów kluczowych w postaci LN.

$LN^+ \rightarrow ?$

$LN^+ \rightarrow \text{LN}$  (zgodnie z prawem zwrotności Armstronga, gdyż  $LN^+ \subseteq LN$ )

$LN^+ \rightarrow \text{LNM}$  (gdyż L determinuje M tj. istnieje zależność funkcyjna  $L \rightarrow M$ )

$LN^+ \rightarrow \text{LNMK}$  (gdyż MN determinuje K tj. istnieje zależność funkcyjna  $MN \rightarrow K$ )

Komentarz:  $NM \rightarrow K$  jest tożsame dla  $MN \rightarrow K$  (kolejność atrybutów nie ma znaczenia)

Ponieważ po prawej stronie pojawiły się wszystkie atrybuty relacji  $R(K, L, M, N)$  to możemy powiedzieć, że atrybuty LN determinują (określa) **wszystkie** atrybuty (K, L, M, N) relacji R. Jest to warunek konieczny, aby uznać atrybuty LN jako **klucz** relacji.

Reasumując otrzymaliśmy trzy klucze MN, KN, KL

Wynik do sprawdzenia tutaj: [http://uisacad5.uis.edu/cgi-bin/mcrem2/database\\_design\\_tool.cgi](http://uisacad5.uis.edu/cgi-bin/mcrem2/database_design_tool.cgi)

### Przykład 3

Dana jest relacja 4-atrybutowa  $R(K, L, M, N)$  dla której zachodzą następujące zależności funkcyjne  $F = \langle K \rightarrow L, L \rightarrow M, M \rightarrow K \rangle$

Dla zależności funkcyjnych z niniejszego przykładu tj.  $F = \langle K \rightarrow L, L \rightarrow M, M \rightarrow K \rangle$  otrzymujemy:

Lewostronne atrybuty kluczowe	Obustronne atrybuty potencjalnie kluczowe	Prawostronne atrybuty niekluczowe
N	MKL	-

Atrybutem kluczowym jest na pewno N, natomiast M, K oraz L są potencjalnymi atrybutami kluczowymi.

Komentarz: Atrybut, który nie znajduje się w żadnej zależności funkcyjnej, ani po lewej, ani po prawej stronie jest zawsze atrybutem kluczowym. Dlatego w tabeli jest wpisywany w kolumnie "Lewostronne".

Dalsze rozważania przebiegają dokładnie tak samo jak w zadaniu **Przykład 2**.

#### Przykład 4

Dana jest relacja 4-atrybutowa **R(K, L, M, N)** dla której zachodzą następujące zależności funkcyjne **F = <K → MN>**

Dla zależności funkcyjnych z niniejszego przykładu tj. **F = <K → MN>** otrzymujemy:

Lewostronne atrybuty kluczowe	Obustronne atrybuty potencjalnie kluczowe	Prawostronne atrybuty niekluczowe
<b>K</b>	-	-

Atrybutem kluczowym jest na pewno **KL** gdyż w przypadku gdy atrybut nie uczestniczy w żadnej zależności funkcyjnej jest na pewno atrybutem kluczowym tj. musi wejść do klucza relacji.