

# Modelowanie baz danych

## 1. Krótki wstęp do relacji

Kluczowym modelem danych jest model relacyjny. W relacyjnym systemie baz danych podstawową strukturą jest **tabela = relacja**. Jedna relacja stanowi zbiór rekordów o identycznej strukturze. Wyobraźmy sobie na przykład relację o nazwie *student*:

### Relacja *student*

Nr indeksu	nazwisko	imię	Rok urodzenia
12345	Kowalczyk	Jan	1983
34567	Nowak	Anna	1985
...	...	...	...

Relacja odpowiada zatem pewnej klasie obiektów świata rzeczywistego, w tym przypadku studentów. Elementy relacji to:

- unikalna nazwa (tutaj: *student*)
- **atrybuty** – kolumny relacji. Nazwa każdego atrybutu znajduje się w nagłówku każdej kolumny. (tutaj: nr indeksu, nazwisko, imię, rok urodzenia)
- **krotki** – wiersze relacji, odpowiadające konkretnym obiektom umieszczonym w relacji. (tutaj: w relacji umieszczono dwa obiekty klasy *student*).

Relacja to **najważniejsza struktura**, na której opierają się dzisiejsze systemy zarządzania bazami danych. Baza danych to zbiór odpowiednio ze sobą powiązanych relacji. Powyższy akapit to jedynie króciutki wstęp mający na celu przybliżenie problemu.

## 2. Projektowanie bazy danych

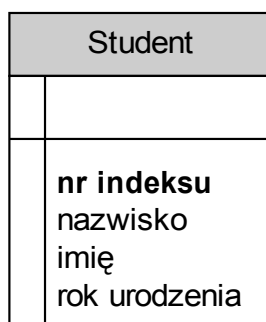
Projektowanie bazy danych odbywa się w następujących etapach:

1. Analiza danych (wybór tych fragmentów rzeczywistości, które mają być modelowane),
2. Ustalenie zależności między wyodrębnionymi składowymi,
3. Zamodelowanie struktury danych. **Struktura to schemat bazy danych, a nie faktyczne dane**. Dopiero po stworzeniu fizycznej bazy danych można umieścić w niej dane. Najbardziej uniwersalnym sposobem opisu struktury bazy danych jest stosowanie **diagramów związków encji**, czyli graficznej reprezentacji danych.
4. Przekształcenie modelu związków encji do postaci relacyjnej.

## 3. Diagramy związków encji - pojęcia

**Encja**. Encja symbolizuje pewien obiekt świata rzeczywistego. Typ encji (stosowany zamiennie z terminem „encja”) to klasa obiektów świata rzeczywistego. Encją, w nawiązaniu do powyższego przykładu może być *student*. Encja jest graficznie reprezentowana w postaci prostokąta. W prostokącie tym muszą być zawarte następujące składowe:

- **nazwa encji**,
- **atrybuty** (w tym identyfikator) + opcjonalnie typy atrybutów.



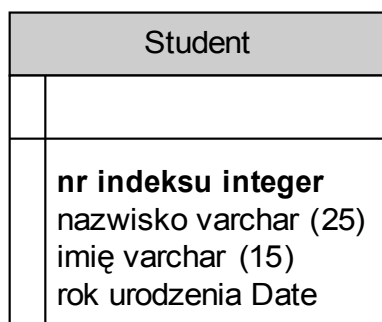
Nazwę encji oddzielamy od pozostałych składowych pogrubioną linią. Poniżej nazwy encji umieszczamy atrybuty, przy czym identyfikator musi być podkreślony.

**Atrybuty.** Jest to zbiór cech, właściwości danego typu encji reprezentującej fragment rzeczywistości. Są one odpowiednikami atrybutów w modelu relacyjnym. Jeden (lub więcej) z atrybutów jest **identyfikatorem** encji – musi być zagwarantowana jego unikalność. W powyższym przykładzie identyfikatorem jest *numer indeksu*. To za jego pomocą możemy jednoznacznie identyfikować obiekty w fizycznej bazie danych (czyli instancje danego typu encji – np. konkretnych *studentów*). W graficznej reprezentacji encji identyfikator jest podkreślonym atrybutem.

Atrybuty mają określone typy. Możliwa (a nawet wskazana) jest notacja, w której w diagramie encji nazwa każdego atrybutu poprzedzona jest nazwą typu (poniżej typy używane w SQL):

- CHAR (n) – tekst o ustalonej długości n znaków,
- VARCHAR (n) – tekst o długości co najwyżej n znaków,
- INT / INTEGER – wartości całkowite,
- FLOAT/REAL oraz DOUBLE PRECISION – liczby zmiennopozycyjne,
- DATE oraz TIME – daty i czas,
- BIT (n) oraz BIT VARYING (n) – ciągi binarne.

Już na etapie modelowania bazy danych warto mieć świadomość typów danych, za pomocą których określimy nasze atrybuty. Encję z przykładu można więc przedstawić bardziej szczegółowo:



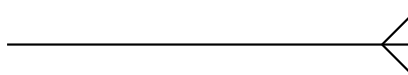
**Związki.** Opisuja bezpośrednie zależności pomiędzy encjami. Parametrami związków są liczebność oraz uczestnictwo (udział) określone za pomocą odpowiednich symboli.

1. **liczebność** związku:

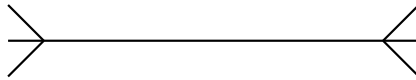
- 1:1 – jeden do jednego:



- 1:N – jeden do wielu:

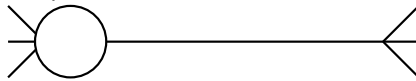


- N:M – wiele do wielu:

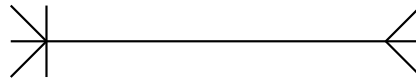


2. **uczestnictwo** (udział):

- minimum 0 (nie musi występować):



- minimum 1 (musi wystąpić):

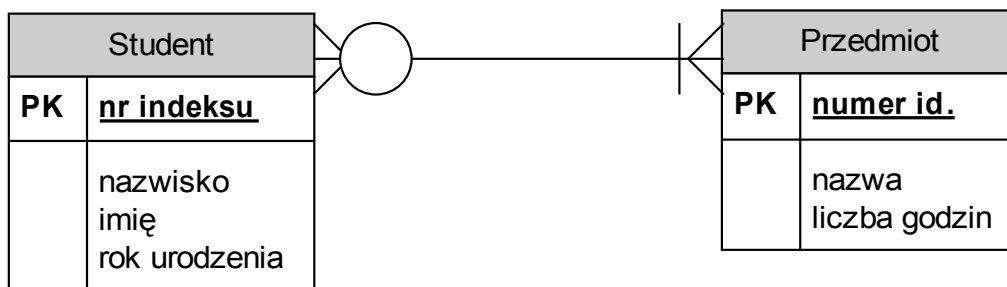


Liczebność oraz uczestnictwo muszą być oznaczone **na jednym rysunku** symbolizującym dany związek. Co więcej, nad linią symbolizującą związek musi znajdować się jego **nazwa**.

Przykład związku:

- **Liczebność.** Dany przedmiot może studiować wielu studentów. Dany student może studiować wiele przedmiotów (związek **wiele do wielu**).
- **Uczestnictwo.** Każdy student **musi** być zapisany na co najmniej jeden przedmiot (udział: minimum 1). Danego przedmiotu **nie musi** studiować żaden student (udział: minimum 0).

Reprezentacja w postaci związku encji:



## 4. Ćwiczenia

1. Zamodeluj małą bazę danych dla dziekanatu, rozbudowując powyższy przykład. Dodaj następujące informacje: wymagania (przedmioty wymagane przed przystąpieniem do danego kursu), wykładowcy, miejsce zajęć.
2. Baza filmów w wypożyczalni DVD zawiera następujące elementy: filmy, gwiazdy, studia filmowe. Zamodeluj diagram związków encji dla tej bazy danych.
  - Dodaj związek, który obrazuje powiązanie danego filmu z jego poprzednimi odcinkami.
  - W bazie należy umieścić informację o gażach, które aktorzy otrzymali za określone filmy. Aktor otrzymuje określoną gażę za dany film. Jej wysokość dla danego aktora jest z kolei zależna od filmu.
  - Studia specjalizują się w określonych gatunkach filmowych.
3. Zamodeluj bazę danych dla następującego problemu:
  - klient firmy handlowej w czasie zakupów nabyć może pewną liczbę produktów,
  - z każdą transakcją wiąże się faktura na określoną kwotę. Na fakturze wyszczególnione są wszystkie nabyte produkty,
  - każda faktura zawiera również informację o pracowniku (lub pracownikach),

którzy brali udział w transakcji,

- każdy towar dostarczony jest do firmy przez co najmniej jednego zewnętrznego dostawcę. Każdy z dostawców może dostarczać pewną grupę produktów.
4. Zamodeluj bazę danych dla następującego problemu:
- W warsztacie samochodowym mogą być naprawiane samochody danych marek.
  - Naprawa wykonywana jest przez określonych pracowników warsztatu,
  - Pracownicy zatrudniani są na zasadzie zleceń, mogą więc świadczyć usługi na rzecz więcej niż jednego warsztatu.
  - Prowadzony jest rejestr usterek dla każdej naprawy,
  - Dla każdego samochodu przechowuje się spis minionych napraw.