

# Modelowanie konceptualne

## Opis rzeczywistości

- klasy, obiekty, relacje
- procesy, zdarzenia

## Pojęcia podstawowe

- **obiekty = encje = instancje** – istniejące niezależnie i dające się zidentyfikować elementy rzeczywistości. Najczęściej opisywane za pomocą rzeczowników lub nazw własnych. Przykłady: **osoba, Jan Kowalski**.
- **atrybuty = własności** – cechy obiektów, posiadające wartość należącą do określonego zbioru (domeny). Najczęściej opisywane za pomocą przymiotników i liczebników. Przykłady: **wysoki, ważący 54kg**.
- **związki = zależności** – elementy rzeczywistości, których istnienie nie jest niezależne – występują pomiędzy dwoma lub większą liczbą obiektów. Opisywane są za pomocą dwu i więcej argumentowych predykatów. Przykłady: **X jest synem Y, X przybył z Y do Z**.
- **klasy = zbiory encji** – zbiory obiektów, posiadających wspólne cechy, należących do jednego rodzaju. Opisywane są najczęściej za pomocą rzeczowników pospolitych w liczbie mnogiej. Przykłady: **ludzie, samochody**.

## Etapy modelowania konceptualnego

- identyfikowanie **obiektów**
- identyfikowanie **atrybutów**
- identyfikowanie **klas obiektów**
- identyfikowanie **zależności** występujących pomiędzy **klasami obiektów**:
  1. **generalizacja/specjalizacja** – klasa bardziej specyficzna, posiada wszystkie cechy klasy bardziej ogólnej. Przykład (k. ogólna – k. specyficzna): **zwierzę – pies, narzędzie – młotek**.
  2. **powiązanie** – relacja występująca pomiędzy obiektami należącymi (najczęściej) do dwóch różnych klas. Przykład: osoba **prowadzi** samochód (tzn. każdy samochód będący w ruchu, jest prowadzony przez jakąś osobę).
  3. **agregacja** – szczególny rodzaj powiązania, w którym można wyróżnić klasę **całość** oraz klasę **część**. Przykład: **biblioteka – książka**.
  4. **kompozycja** – szczególny rodzaj agregacji, w którym obiekty-części nie mogą występować bez obiektu-całości. Przykład: **ciało – organy, budynek – pomieszczenia**.

Jaka jest zasadnicza różnica między **agregacją** a **kompozycją**?
- krotność powiązania – dopuszczalna liczba powiązań występujących pomiędzy obiektami dwóch klas:
  1. **1 do 1 = jeden do jeden** – jeden obiekt pierwszej klasy może być powiązany tylko i wyłącznie z jednym obiektem drugiej klasy.

Przykład: **samochód – kierownica**.

2. **1 do n = jeden do wiele** – jeden obiekt pierwszej klasy może być powiązany z jednym lub więcej obiektami drugiej klasy, ale jeden obiekt drugiej klasy może być powiązany tylko z jednym obiektem klasy pierwszej.

Przykład: **żaglówka – członkowie załogi**.

3. **n do n = wiele do wielu** – zarówno obiekt klasy pierwszej może być powiązany z jednym lub więcej obiektami klasy drugiej, jak i obiekt klasy drugiej może być powiązany z jednym lub więcej obiektami klasy pierwszej.

Przykład: **aktor – sztuka**.

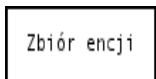
Powyższe przypadki nie wyczerpują wszystkich możliwości, ale reprezentują najczęściej spotykane zależności pomiędzy klasami.

# Diagramy ER


Diagramy ER = diagramy związków-encji = entity-relationship diagrams

## Elementy podstawowe

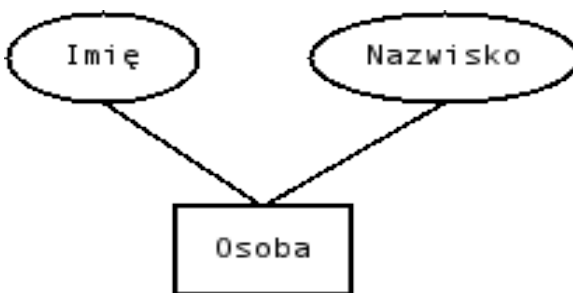
Diagramy związków-encji pozwalają w sposób graficzny opisywać model konceptualny. Do podstawowych elementów diagramów ER należą:

 **zbiory encji** reprezentowane przez prostokąty, które w środku zawierają nazwę zbioru

 **atrybuty** reprezentowane przez elipsy, które w środku zawierają nazwę atrybutu

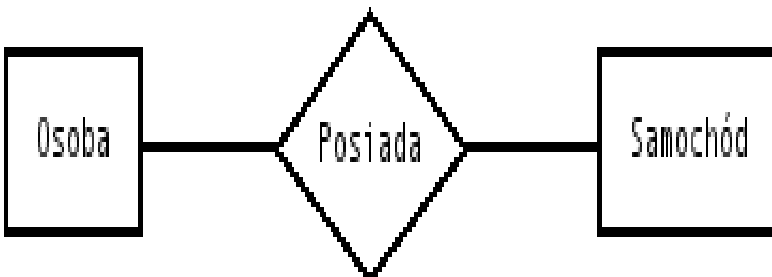
 **związki** reprezentowane przez równoległoboki, które w środku zawierają nazwę związku

Zbiory encji opisywane są za pomocą atrybutów, które łączy się z nimi za pomocą pojedynczej linii prostej:

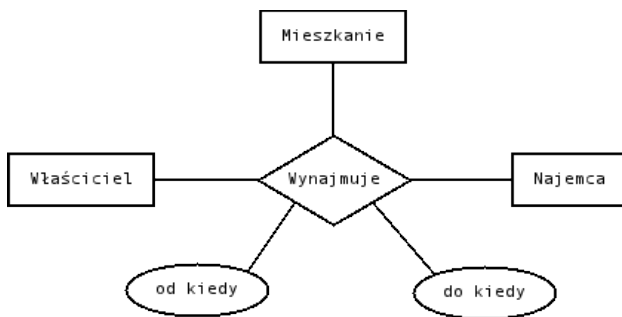


Uwaga: wewnątrz elipsy wprowadzana jest **nazwa atrybutu**, a nie jego **wartość**. Zatem niepoprawnym atrybutem jest np. **czerwony**, gdyż jest to wartość atrybutu, który powinien nazywać się **kolor**.

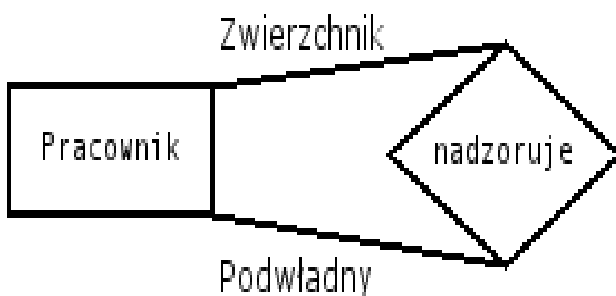
Związki zachodzące pomiędzy zbiorami encji przedstawia się w postaci równoległoboku, z którego wychodzą linie proste do wszystkich zbiorów encji, należących do danego związku:



Związki zachodzące pomiędzy zbiorami encji mogą być więcej niż dwuargumentowe, mogą również posiadać atrybuty:



Związki mogą zachodzić pomiędzy encjami należącymi do jednego zbioru. Taka sytuacja wymaga jednak opisanie ról w jakich występują poszczególne encje:



### Krotność związków

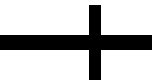

Krotność związków zbiorów encji jest cechą, która opisuje *z iloma encjami należącymi do jednego zbioru encji może być połączona jedna encja należąca do drugiego zbioru encji*.

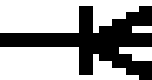

Jak wskazaliśmy wcześniej (w odniesieniu do obiektów i klas), istnieją trzy główne typy związków:

- 1 do 1
- 1 do n
- n do n

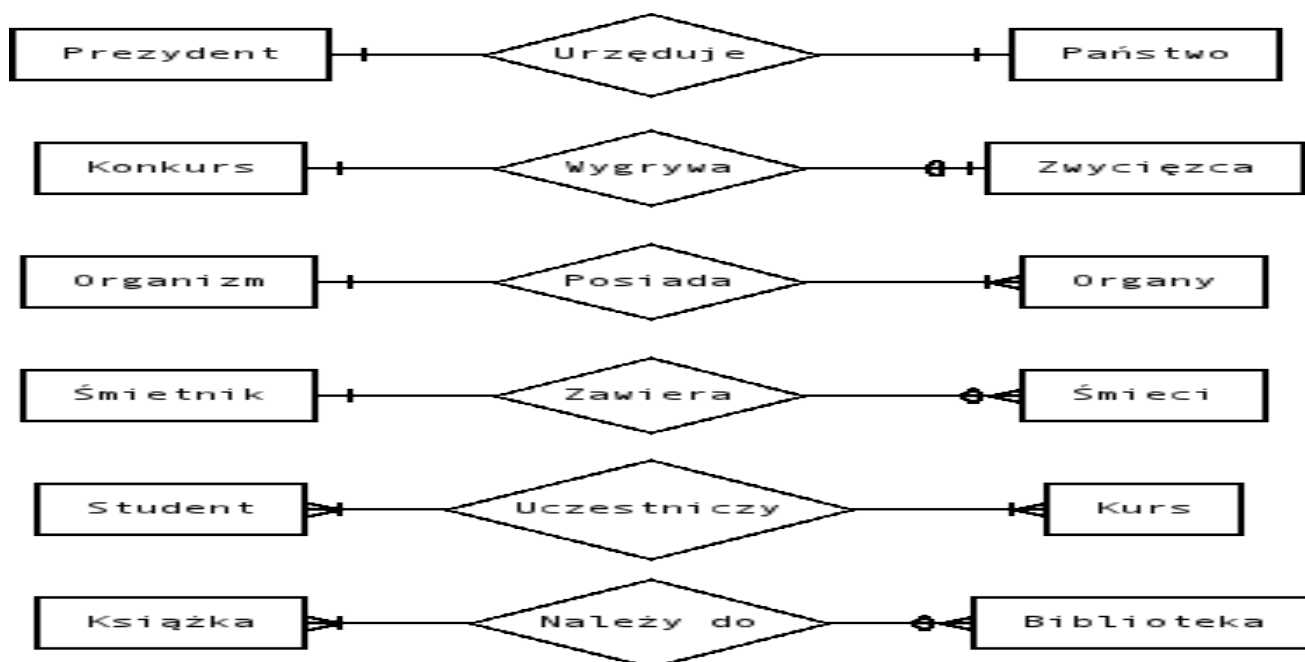
Tym niemniej wykorzystując tylko te trzy typy, nie jesteśmy w stanie wystarczająco precyzyjnie opisać przypadków z jakimi spotykamy się w rzeczywistości. Np. często zachodzi potrzeba odróżnienia sytuacji, w której jedna encja jednego zbioru **może** być związana z jedną encją drugiego zbioru, od sytuacji, w której relacja ta jest obligatoryjna (tzn. jedna encja pierwszego zbioru **musi** być związana z jedną encją drugiego zbioru).

W literaturze poświęconej modelowaniu baz danych można spotkać się z symboliką, która oparta jest na modelu, w którym rozróżnia się tylko trzy podstawowe przypadki. Tym niemniej w praktyce wykorzystywane są notacje, które pozwalają precyzyjnie opisywać te zależności. Niestety w odniesieniu do diagramów ERD nie wypracowano jednego standardu oznaczania krotności (jak ma to miejsce w przypadku standardu modelowania obiektowego [UML](#)), aczkolwiek najczęściej można spotkać się z następującą notacją (rozpowszechnioną m. in. w dokumentacji Oracle):

-  – powiązanie jednokrotne, obowiązkowe (tylko jeden)
-  – powiązanie jednokrotne, opcjonalne (jeden lub zero)

-  – powiązanie wielokrotne, obowiązkowe (co najmniej jeden)
-  – powiązanie wielokrotne, opcjonalne (zero lub więcej)

Poniżej zaś przedstawiamy przykład, który dobrze ilustruje różne kombinacje krotności:



Uwaga: symbole krotności interpretowane są zawsze w ten sposób, że *jedna encja* z jednego zbioru encji może być powiązana z taką liczbą encji należących do *drugiego zbioru*, na jaką wskazuje symbol leżący bezpośrednio przy *drugim zbiorze encji*.

W powyższym przykładzie – jeden organizm może posiadać jedno lub więcej organów, natomiast jeden organ musi należeć do dokładnie jednego organizmu.

## Typy atrybutów

W modelu ER wyróżnić można trzy specjalne typy atrybutów:

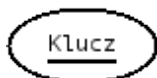
- kluczowe
- pochodne
- częściowe kluczowe

Jednym z ważniejszych etapów projektowania modelu koncepcyjnego, jest zidentyfikowanie **atrybutów kluczowych**, które tworzą *klucz* encji. *Atrybuty kluczowe* to atrybuty, które w sposób jednoznaczny identyfikują encję, która je posiada. Znaczy to, że nie mogą występować dwie encje, które posiadają taką samą wartość atrybutów kluczowych.

Klucz może być pojedynczym atrybutem lub zbiorem atrybutów. W drugim przypadku wymóg jednoznaczności odnosi się do zbioru jako całości, a nie poszczególnych atrybutów, tzn. wymóg jednoznaczności odnosi się do kombinacji wartości, a nie każdej wartości z osobna. Np. klucze (1, 2) oraz (1, 3) będą identyfikowały odmienne encje, pomimo tego, że wartość pierwszego atrybutu jest w obu przypadkach taka sama.

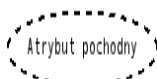
Encja może posiadać wiele kluczy. W takiej sytuacji jeden z nich jest wyróżniony i nazywa się go **kluczem głównym**. Pozostałe klucze, to **klucze alternatywne**. Ponadto **każdy zbiór encji musi posiadać przynajmniej jeden klucz**.

Atrybuty kluczowe wyróżnia się poprzez podkreślenie ich nazwy:



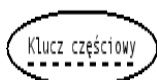
Innym ważnym typem atrybutów, są **atrybuty pochodne**. Ten typ atrybutów wykorzystywany jest do reprezentowania informacji, które zawsze mogą być obliczone na podstawie danych zgromadzonych w bazie, ale zachodzi potrzeba reprezentowania ich w modelu. Przykładem atrybutu tego rodzaju może być np. średnia ocena publikowanego materiału, wyliczana na podstawie głosów oddanych przez czytelników.

Atrybuty pochodne reprezentowane są za pomocą owalu, którego brzeg rysowany jest linią przerywaną:



W diagramach ERD wykorzystuje się również **klucze częściowe**. Klucze częściowe pełnią tę samą rolę, co zwykłe kluczowe, z tą różnicą, że samodzielnie nie identyfikują encji. Partycypują w kluczach, które tworzone są z atrybutów innych encji i występują wyłącznie w *słabych encjach*.

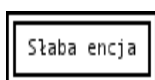
Klucze częściowe wyróżniane są za pomocą podkreślenia linią przerywaną:



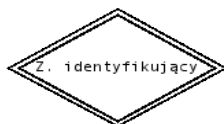
## Słabe encje

**Słabe encje** to szczególny typ encji, których klucze składają się (przynajmniej częściowo) z atrybutów kluczowych innych encji. Jeśli encja jest słaba, to pomiędzy zbiorem encji, których klucze wykorzystywane są w jej kluczu, a zbiorem do którego on należy, musi zachodzić związek, którego krotność po stronie przeciwnej słabej encji, wynosi dokładnie jeden. W przeciwnym bowiem razie, mogłaby wystąpić sytuacja, w której słaba encja nie posiadałaby jednoznacznego identyfikatora.

**Słabe encje** reprezentowane są za pomocą prostokątów z podwójną ramką:



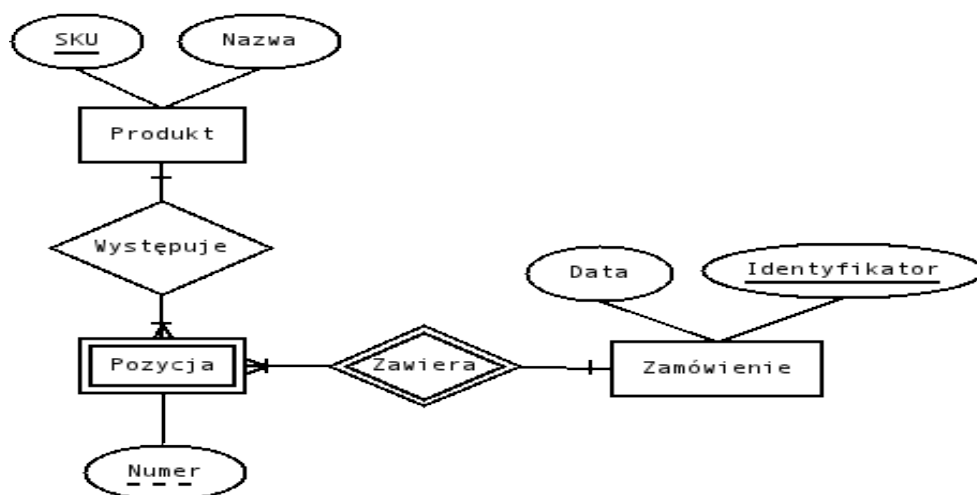
**Związki identyfikujące**, które wskazują zbiory encji, których klucze wykorzystywane są przez zbiory słabych encji, reprezentowane są za pomocą podwójnego równoległoboku:



Słabe encje wykorzystywane są zwykle wtedy, gdy mamy do czynienia z kompozycją lub związkiem wiele do wiele przekształconym w zbiór encji, tzn. słabe encje zawsze występują jako elementy pewnej struktury wyższego rzędu.

Jako przykład można podać zamówienie, które posiada swój unikalny identyfikator oraz pozycje, których porządek jest ustalony przez użytkownika. Zamiast przypisywać każdej pozycji odrębny identyfikator (klucz) i dodatkowo numer porządkowy na zamówieniu, prościej jest opisać pozycję, za pomocą klucza składającego się z identyfikatora zamówienia i numeru pozycji.

Sytuacja ta przedstawiona jest na poniższym diagramie:



## Relacja “is\_a”

Relacja *is\_a* (“X is a Y”), występuje pomiędzy zbiorami encji, spośród których jeden może być określony jako specjalizacja drugiego. Tzn. każda encja należąca do pierwszego zbioru posiada wszystkie cechy encji należących do zbioru drugiego, ale ponadto posiada pewne cechy specyficzne. Przez cechy rozumie się tutaj nie tylko atrybuty, ale również wszystkie związki w jakie wchodzi dany zbiór encji.

Na diagramach ERD relacja *is\_a* oznaczana jest za pomocą trójkąta, wewnątrz którego znajduje się napis “Isa”:



Relacja *is\_a* jest jedną z podstawowych relacji występujących pomiędzy pojęciami. Typowy zwrot, który służy do jej wyrażania, to “X jest rodzajem Y” lub “X jest Y”, gdzie X i Y są pewnymi pojęciami ogólnymi (np. *wilk jest zwierzęciem*, *lazurowy jest rodzajem niebieskiego*, *van jest rodzajem samochodu*).

Klasyczna **definicja arystotelesowska** składała się z dwóch elementów: *genus proximum* (rodzaj najbliższy) i *differentia specifica* (różnica gatunkowa). Jest ona jednym z podstawowych sposobów

definiowania słów, a w konsekwencji organizowania wiedzy. Widać wyraźnie, że jest ona oparta o relację “is\_a” - aby zdefiniować słowo należy wskazać inne, którego znaczenie jest bardziej ogólne od słowa definiowanego.

Poniżej przedstawiony jest **przykład** wykorzystujący relacji “is\_a”. Został w nim wyabstrahowany jeden zbiór encji: *ciała niebieskie*. Wszystkie ciała niebieskie posiadają masę, temperaturę i promień (można przyjąć dla uproszczenia, że zajmujemy się tylko ciałami kulistymi). Niektóre spośród nich posiadają jednak cechy specyficzne – gwiazdom i planetom możemy przypisać jasność, a planetom i księżycom – promień orbity, po której się poruszają. Ponadto istotny jest również fakt, że księżyce krążą wokół planet, a planety wokół gwiazd. Wszystkie te informacje zostały uwzględnione na diagramie.

