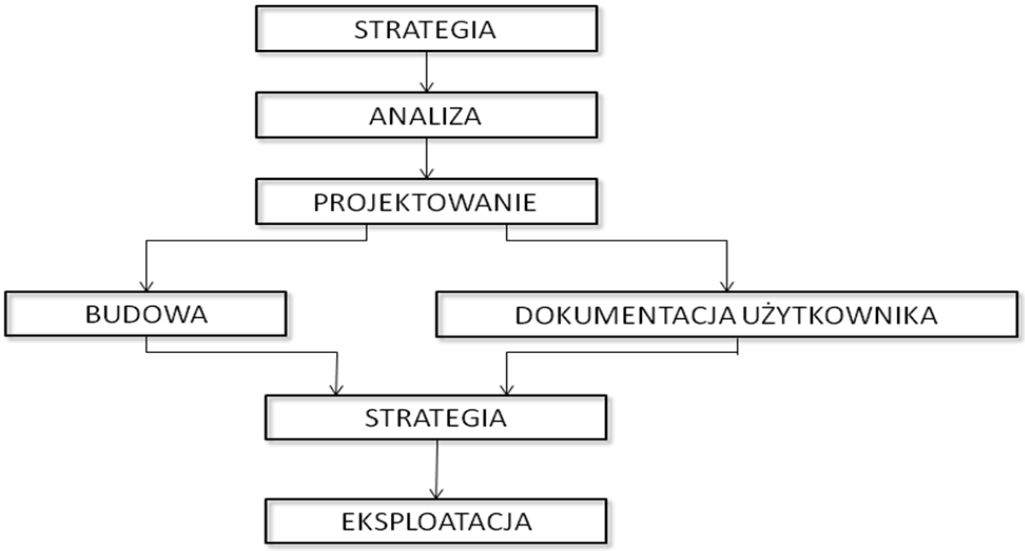
Modelowanie związków encji

1. Cel modelowania - tworzenia związków encji

Metoda modelowania – tworzenie związków encji (ERD) – odnosi się do etapów strategii i analizy cyklu życia systemu informacyjnego.



Cykl życia systemu informacyjnego

Po zakończeniu fazy gromadzenia i analizy informacji oraz wymagań od przyszłych użytkowników bazy danych, tj. po uzyskaniu informacji o wymaganiach użytkowników dotyczących przyszłej aplikacji bazy danych, oraz po ich odpowiednim udokumentowaniu, możemy przystąpić do fazy projektowania.

Jednym z trudniejszych aspektów projektowania bazy danych jest fakt, że projektanci, programiści i użytkownicy maja tendencję do odmiennego widzenia danych i sposobów ich wykorzystania. W ogóle to wydaje się, iż samo nasze postrzeganie świata jest samym w sobie złożonym modelem, który tworzymy w naszych umysłach na bazie gromadzonej informacji, Modelu tego używamy w chwili, gdy napotkamy nowe sytuacje, aby je wyjaśnić i dać sobie radę z powstającymi problemami. Nasze wyobrażenia i twórcze myślenie wynikają z naszego intelektualnego modelu świata.

Musi zatem istnieć pewien model, który musi być jednoznaczny, a przy tym mało techniczny, który byłby zrozumiały dla wszystkich, którzy uczestniczą w tworzeniu bazy. Jedną z metod pozwalającą uzyskać taki model jest modelowanie związków encji.

2. Określenie i przeznaczenie modelowania związków encji

Modelowanie związków encji:

- jest metodą określania potrzeb informacyjnych firmy lub organizacji,
- stanowi trwałą podstawę do budowania odpowiednich, wysokiej jakości systemów bazodanowych, które będą realizowały potrzeby danej firmy lub organizacji.
- reprezentuje podejście zstępujące dotyczące projektowania bazy danych tzn. zaczyna się od wyodrębnienia istotnych danych, zwanych *encjami*, oraz tych *związków* pomiędzy nimi, które powinny być reprezentowane w modelu. Następnie do modelu dodawanych jest coraz więcej szczegółów, takich jak informacje o encjach i związkach, które powinny być przechowywane w bazie – zwane *atrybutami*, a także *więzy* (warunki) dotyczące encji, związków i atrybutów.

- ❖ jest techniką, którą musi opanować każdy projektant bazy danych.

Modelowanie związków encji ma na celu:

- ➢ dostarczenie dokładnego modelu potrzeb informacyjnych przedsiębiorstwa, który stanowiłby podstawę do konstruowania nowych lub ulepszonych systemów,
- ➢ dostarczenie modelu niezależnego od sposobu przechowywania danych i od metod dostępu do nich, umożliwiającego podejmowanie celowych decyzji, jeśli chodzi o metod implementacyjne i współdziałanie z istniejącymi systemami.

3. Struktura diagramu związków encji

Model związków encji zwany jest też modelem ER (ang. *Entity – Relationship Model*).

Diagram związków encji, zwany jest też *diagramem ERD* (ang. *Entity – Relationship Diagram*)

Model danych w postaci *diagramu związków encji* zawiera trzy rodzaje elementów:

- encje,
- atrybuty,
- związki.

4. Encja i jej graficzna reprezentacja

Encja (obiekt) – jest rzeczą lub obiektem, rzeczywistym lub wyobrażonym, czymś co ma znaczenie dla nas i jest odróżnialne od innych rzeczy i obiektów, czymś o czym informacje powinny być znane i przechowywane.

Reprezentacją graficzną encji jest ramka (inaczej prostokąt o zaokrąglonych rogach) z umieszczoną w środku nazwą encji. Nazwa powinna być pisana w liczbie pojedynczej i napisana wielkimi literami. Uwaga. Jeśli w danym kontekście istnieją różne słowa będące nazwami tej samej encji, można używać synonimów. Jedna nazwę wybieramy jako główną; natomiast synonimy (wszystkie pisane wielkimi literami) wymieniamy, poprzedzając je ukośną kreską.

Ramka:

- może mieć dowolny rozmiar i kształt,
- powinna być na tyle duża, aby pomieścić jednoznaczną nazwę (bez żadnych skrótów) i aby rysowanie diagramu związków encji było wygodne,
- niekiedy powinna być bardziej rozciągnięta, aby umożliwić łatwiejsze doprowadzenie do niej większej liczby linii związków bez niepotrzebnego przecinania się linii.

5. Własności encji

Nazwa encji musi dokładnie reprezentować typ lub klasę rzeczy – a nie jakąkolwiek jej instancję (tj. wystąpienie). Np. nazwy kin „Bałtyk", czy „Silver-Screen" nie mogą być użyte jako nazwy encji, gdyż encją jest KINO, zaś wymienione nazwy kin są tylko instancjami tej encji.

Reguły dotyczące encji:

- ❖ wszystkie encje się wzajemnie wykluczają, tzn. iż każdy przedmiot lub rzecz może być reprezentowany tylko przez jedną encję,
- ❖ każda encja musi być jednoznacznie identyfikowalna, tzn. że każda instancja (wystąpienie) encji musi być wyraźnie odróżnialna od wszystkich innych instancji tego typu encji.

6. Związki i ich graficzna reprezentacja

Związki – to zbiór istotnych powiązań pomiędzy zbiorami encjami. A zatem *związek* to uporządkowana lista encji, przy czym poszczególne encje mogą występować wielokrotnie.

Inaczej mówiąc, każdy *związek* określa pewną relację między zbiorami egzemplarzy encji wchodzącymi w skład związku. W szczególnym przypadku może być powiązaniem tej samej encji ze sobą. Relację te nazywamy *instancją związku.*

$$Z(E_1,\ldots E_n)\ ,$$

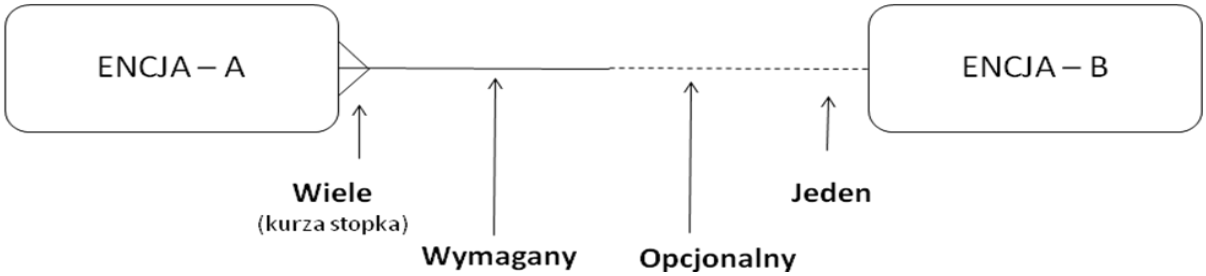
gdzie $E_1,\ldots E_n$ – encje występujące w związku Z.

Reprezentacją graficzną związku jest linia (krzywa) łącząca ze sobą dwie ramki encji lub rekurencyjnie jedną ramkę ze sobą.

Każdy związek ma:

- przypisaną nazwę,
- stopień,
- opcjonalność (opcjonalny czy wymagany).

7. Związki i ich graficzna reprezentacja

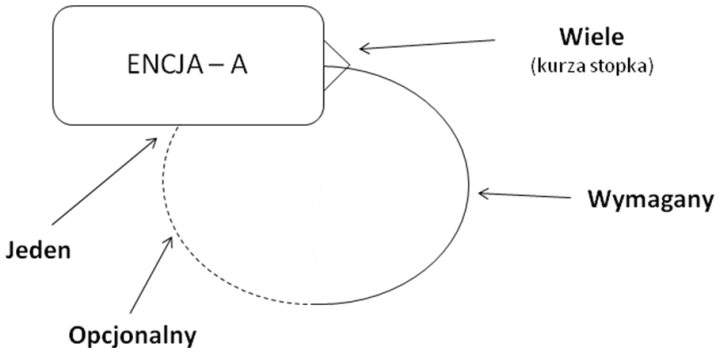


Przy liczebności „wiele" linia związku dotyka ramki przedstawiającej encję w trzech punktach. Połączenie to nosi nazwę „kurzej stopki" (ang. *crowsfoot*). Przy liczebności związku „jeden" linia związku dotyka ramki tylko w jednym punkcie.

Gdy związek jest wymagany, wtedy reprezentująca go połowa linii jest ciągła, natomiast gdy jest opcjonalny, to odpowiadający koniec linii rysujemy przerywany.

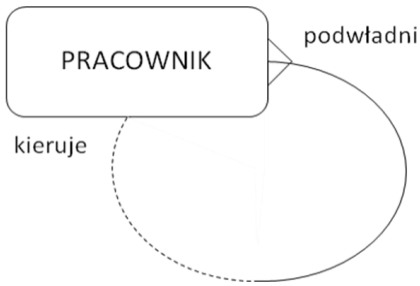
8. Związek rekurencyjny i jego reprezentacja graficzna

Związek rekurencyjny – to związek, w którym ta sama encja występuje więcej niż raz w różnych rolach.



9. Przykład nazwanego związku rekurencyjnego

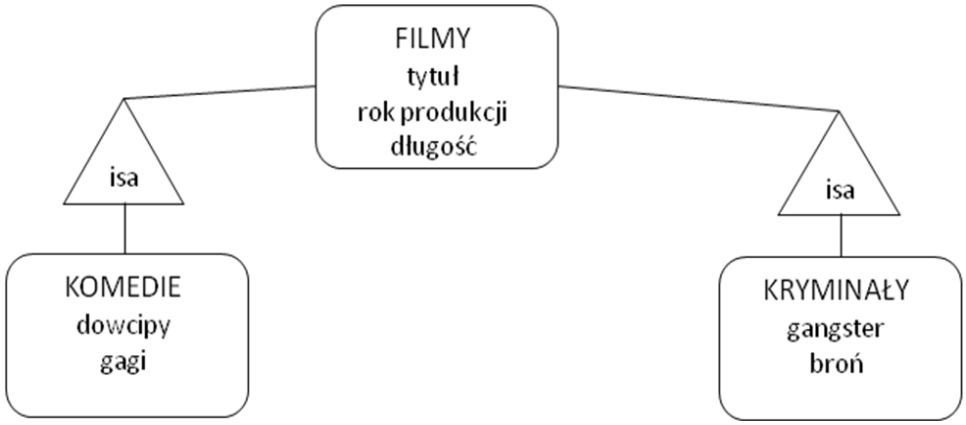
PRZYKŁAD. Rozważmy następujący przykład związku rekurencyjnego.



Powyższy diagram możemy przeczytać w następujący sposób:
Każdy PRACOWNIK będący KIEROWNIKIEM może kierować pracą jednego lub więcej PRACOWNIKÓW.
oraz
Każdy PRACOWNIK musi mieć jednego i tylko jednego KIEROWNIKA (tj. PRACOWNIKA, który jest jego bezpośrednim przełożonym).

10. Związek typu „isa" i podklasy

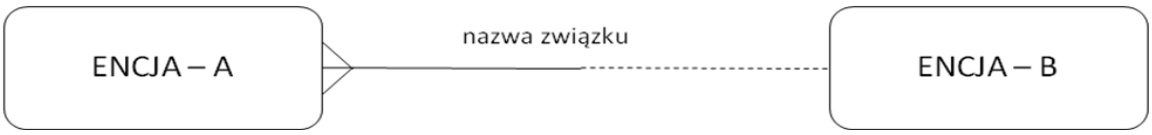
Zbiór encji często zawiera pewne encje o specjalnych własnościach, których nie mają wszystkie elementy tego zbioru. W takim przypadku użyteczne jest zdefiniowanie specjalnych zbiorów encji, tzw. **podklas**, które zawierają te dodatkowe atrybuty i/lub związki. Zbiór encji łączymy z jego podklasami przy wykorzystaniu **związku „isa"** (tzn. stwierdzenie „A isa B" oznacza związek ze zbioru encji A do zbioru encji B).
Związek isa jest specjalnym rodzajem związku, i aby podkreślić jego wyjątkowość, używa się dla niego specjalnego oznaczenia – każdy związek isa reprezentuje się za pomocą trójkąta. Jedna krawędź trójkąta jest podłączona do podklasy, a przeciwległy wierzchołek jest połączony z nadklasą. Każdy związek isa jest typu jeden-do-jednego.
PRZYKŁAD. Wśród filmów występują kryminały, komedie, przygodowe, kreskówki, itd. Dla poszczególnych typów można zdefiniować odpowiednią podklasę encji *Filmy*, np. wybierzmy podklasy *Kryminały* i *Komedie*.



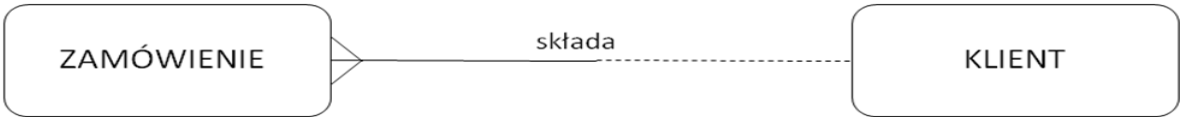
Podklasy i związki typu isa na diagramie ER

11. Nazywanie związków

Każdy związek posiada swoją *nazwę*, która zapisana małymi literami, jest umieszczana przy linii reprezentującej graficznie ten związek.



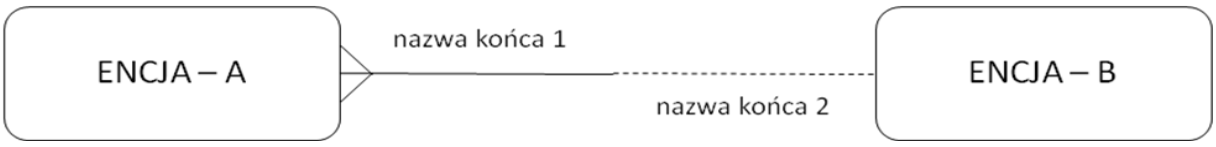
PRZYKŁAD. Przykład nazwanego związku KLIENT *składa* ZAMÓWIENIE.



Powyższy diagram możemy przeczytać w następujący sposób:
Każde ZAMÓWIENIE jest **składane** przez jednego i tylko jednego klienta.
oraz
Każdy KLIENT może **złożyć** jedno lub więcej ZAMÓWIEŃ.

12. Inna konwencja nazewnictwa związków

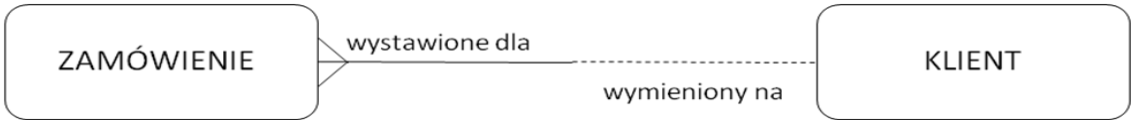
W innej konwencji nazwy mogą być także przypisane także końcom linii związku - wtedy nazwy te określają, jakie role odgrywają poszczególne zbiory encji w danym związku i dlatego z tego powodu bywają także określane mianem *nazw ról*.



W przypadku, gdy dany koniec jest wymagany, używa się wyrażenia „musi być" (czasami „musi mieć") przed nazwą danego końca związku. Tymczasem dla nazw opcjonalnych końców używa się wyrażenia „może być" (czasami „może mieć").
A zatem powyżej zaprezentowany diagram należy odczytywać od strony lewej do prawej w następujący sposób:
Każda ENCJA-A musi być nazwa-końca-1 jedna i tylko jedna ENCJA-B.
Natomiast czytając od strony prawej do lewej, otrzymujemy:
Każda ENCJA-B może być nazwa-końca-2 jedna lub więcej ENCJA-A.

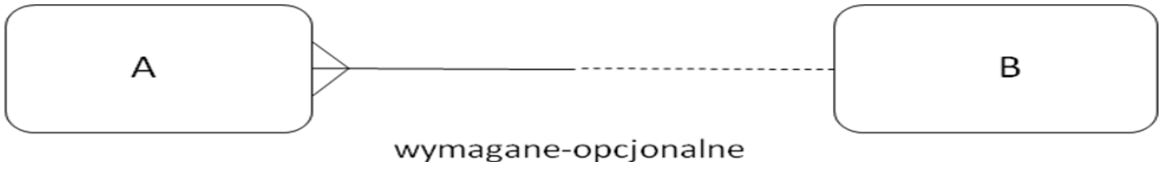
13. Nazywanie związków

PRZYKŁAD. Przykład nazwanego związku.



Powyższy diagram możemy przeczytać w następujący sposób:
Każde ZAMÓWIENIE musi być **wystawione dla** jednego i tylko jednego klienta.
oraz
Każdy KLIENT może być w **wymieniony na** jednym lub więcej ZAMÓWIENIACH.
UWAGA. Nazwy encji w liczbie mnogiej używa się w przypadku, gdy liczebność (stopień) jest określony jako wiele, np. ZAMÓWIENIACH. Liczebność wiele jest odczytywana jako jeden lub więcej, natomiast liczebność jeden – jako jeden i tylko jeden.

14. Wymagalność i opcjonalność związku



Każda instancja encji A może istnieć tylko w ramach nazwanego kontekstu jednej i tylko jednej encji B. Z drugiej strony, encje B mogą istnieć niezależnie od tego, czy są powiązane z jakimś A.
Jako przykład można rozważyć jeszcze raz KLIENTA, który może zostać zapisany w katalogu klientów, choć nie złożył jeszcze ani jednego zamówienia.

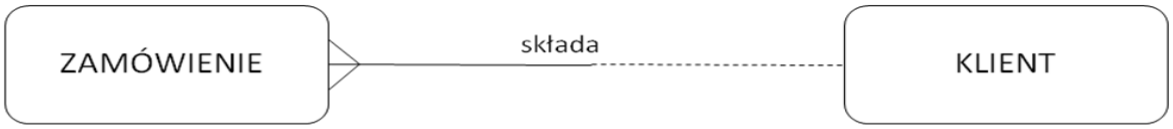
wymagane-wymagane

Nie można utworzyć instancji encji B bez jednoczesnego utworzenia co najmniej jednego związanego z nią A.
W rozważanym później przykładzie nie ma REJSU samolotu, jeśli nie ma w jego ramach choć jednego ODLOTU.

15. Stopień związku

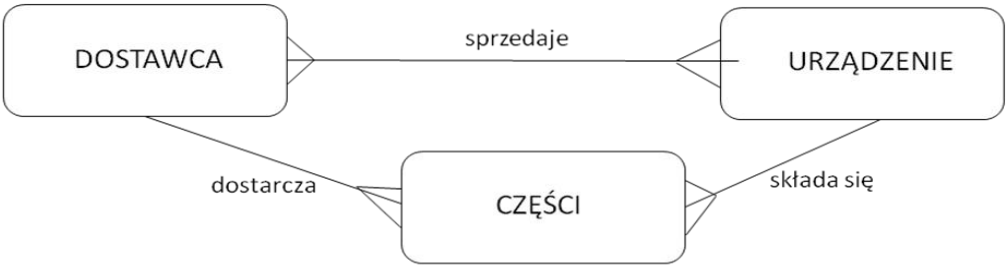
Stopień związku jest określony liczbą uczestniczących w nim encji.
Wyróżniamy:

- związki dwuargumentowe tzw. związki binarne – związki o stopniu dwa



Przykład związku binarnego: KLIENT *składa* ZAMÓWIENIE

- związki potrójne (trójargumentowe) tzw. ternarne – związki o stopniu trzy



Przykład związku ternarnego: DOSTAWCA sprzedaje URZĄDZENIE.
DOSTAWCA *dostarcza* CZĘŚCI *do* URZĄDZENIA.
URZĄDZENIE *składa się* z CZĘŚCI.

16. Rodzaje związków binarnych

- związek wieloznaczny wiele-do-wiele
  np. Każdy ZAMÓWIENIE może zawierać wiele PRODUKTÓW.
  Każdy PRODUKT może być sprzedany na wielu ZAMÓWIENIACH.,
- związek jednoznaczny jeden-do-wiele, wiele-do-jeden
  np. Każda DZIAŁ może zatrudniać wiele OSÓB
  Każda OSOBA może być zatrudniona w co najwyżej jednym DZIALE.
- związek jednoznaczny jeden-do-jeden
  np. Każdy RODZIC jest OSOBĄ
  Tylko OSOBA może być RODZICEM.

Specyficznym rodzajem związków binarnych są już omówione związki rekurencyjne.

17. Efektywne implementacje związków

Tylko w przypadku związków binarnych istnieją ich efektywne implementacje.
W przypadku związków składających się z więcej niż dwóch encji, istnieje metoda przekształcająca je do związków binarnych. Pozwala to otrzymać diagram związków encji zawierający tylko związki binarne związki jednoznaczne.
W tym celu należy:

- dla każdego związku o liczbie encji większej niż dwa

$$Z(E_1,\ldots E_n)\ ,\ n > 2,$$

wprowadzamy nową encję E_0 i n jednoznacznych związków binarnych $Z(E_0,E_i)$ łączących nową encję ze starymi. Klucz encji E_0 jest (przynajmniej w początkowym okresie) sumą kluczy encji $E_1,\ldots E_n$.

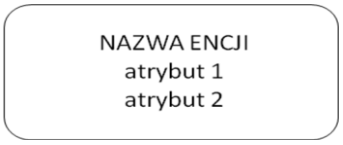
- dla każdego wieloznacznego związku binarnego $Z(E_1,E_2)$ wprowadzamy nową encję E_0 i dwa związki jednoznaczne $Z_1(E_0,E_1)$, $Z_2(E_0,E_2)$ łączące nowy zbiór encji ze starymi. Klucz encji E_0 (przynajmniej w początkowym okresie modelowania) jest sumą kluczy encji E_1 i E_2.

18. Atrybut i jego graficzna reprezentacja

Atrybut – jest to własność encji danego typu, mająca znaczenie dla encji reprezentowana pewną wartością, np. liczbą rzeczywistą, obrazem, napisem, zapachem, itp.

Inaczej mówiąc, *atrybut* – jest dowolnym szczegółem służącym do kwalifikowania, identyfikowania, klasyfikowania, określania ilości lub wyrażania stanu encji.

Aby w sposób graficzny przedstawić atrybut w diagramie, należy w ramkę encji wpisać małymi literami jego nazwę w liczbie pojedynczej. Opcjonalnie można podać jego wartości.



UWAGA. Nie jest koniecznym umieszczanie atrybutów na diagramach związków encji, ale dodanie jednego czy dwóch podczas tworzenia diagramu ułatwia jego tworzenie.

19. Rodzaje atrybutów

- atrybut jednowartościowy – przyjmuje tylko jedną wartość dla każdego wystąpienia encji, np. dla każdego wystąpienia encji KINO atrybut *kodkina* ma pojedynczą wartość np. SC001 i z tego powodu atrybut *kodkina* jest jednowartościowy.
- atrybut wielowartościowy – może przyjmować wiele wartości dla pojedynczego wystąpienia encji, np. dla każdego z wystąpień encji KLIENT atrybut *telefonklienta* może zawierać wiele wartości (gdyż klient może posiadać więcej niż jeden telefon). Atrybut wielowartościowy może zawierać zbiór o ograniczonej z dołu i z góry liczbie elementów (np. atrybut *telefonklienta* może zawierać minimalnie jeden, a maksymalnie trzy numery telefonów.
- atrybut pochodny – to atrybut reprezentujący wartość, która jest wyliczana z podobnego atrybutu lub ze zbioru atrybutów, niekoniecznie pochodzącego z tego samego zbioru encji, np. czas realizacji zamówienia (liczba dni) , jaki upłynął od złożenia zamówienia przez klienta do momentu wysyłki zamówienia może być wyliczony jako różnica tych dat.

20. Klucze

Klucz kandydujący – to najmniejszy zbiór atrybutów jednoznacznie identyfikujący każde wystąpienie encji w zbiorze encji.

Każdy zbiór encji może mieć więcej niż jeden klucz kandydujący.

Klucze kandydujące dzielą się na:

- klucze proste – składające się z dokładnie jednego atrybutu,

- klucze złożone – składające się z więcej niż jednego atrybutu.

Klucz główny – to klucz kandydujący, który został wybrany do jednoznacznej identyfikacji każdego z wystąpień encji w zbiorze encji. Każdy zbiór encji może mieć dokładnie tylko jeden klucz główny.

21. Zbiory słabych encji

W pewnych przypadkach, niektóre, a nawet wszystkie atrybuty klucza zbioru encji wybiera się z innego zbioru encji. Taki zbiór encji określa się mianem **zbioru słabych encji**.

Kolekcja wszystkich podobnych encji tworzy zbiór encji.

Przykład. Każdy wydział pewnej uczelni wyższej może mieć więcej niż jeden kierunek w swojej ofercie edukacyjnej. Wewnątrz danej uczelni kierunki mogą być oznaczone numerami kierunek1, kierunek2, itd. Ale inny wydział tej uczelni może używać tego samego sposobu oznaczania kierunków, a zatem atrybut *nr_kierunek* nie jest kluczem dla kierunku. Aby jednoznacznie określić kierunek, trzeba określić zarówno nazwę wydziału, jak i numer kierunku, który jest w jej ofercie.

Zbiór słabych encji dla kierunków i jego powiązania

Aby jednoznacznie określić kierunek danego wydziału dla tej uczelni, trzeba podać zarówno nazwę wydziału, do którego należy kierunek, jak i numer kierunku. A zatem klucz w zbiorze słabych encji *Kierunki* składa się z jego własnego atrybutu *nr_kierunek*, a także atrybutu *nazwa_wydziału,* należącego do encji *Wydziały*, a określającego jednoznacznie wydział w związku typu wiele-do-jeden o nazwie *należy do* z danym kierunkiem.

Zbiory słabych encji oznaczamy ramką o podwójnych krawędziach, zaś związki wspierające typu wiele-do-jeden rombem o podwójnych krawędziach.

Następująca zasada jest podsumowaniem konwencji przedstawionej na poprzednim slajdzie:

Jeśli jakiś zbiór encji E jest oznaczony podwójną ramką, to musi być słaby. Jego jednoznaczny klucz składa się zatem z tych jego atrybutów, które są podkreślone (może takich elementów wcale nie być) oraz z atrybutów klucza tych encji, z którymi encja E jest połączona poprzez związki typu wiele-do-jeden, a które zostały oznaczone podwójnymi krawędziami.

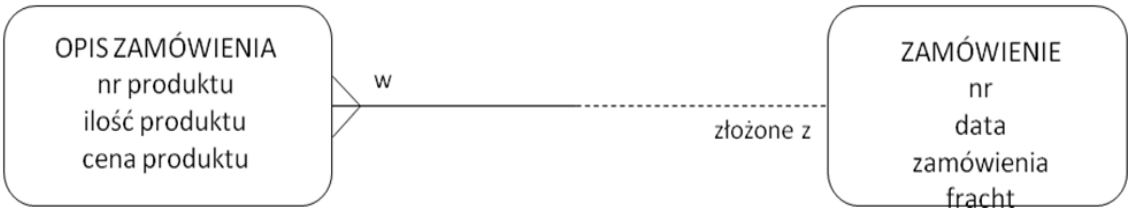
22. Zasady dotyczące atrybutów

- Atrybut musi opisywać encję, przy której jest wyszczególniony – należy przestrzegać zasady, aby nie przypisywać atrybutów encjom, których one nie opisują.
- Nie należy podawać nazw encji jako części nazwy atrybutu – jest to zbyteczne, ponieważ atrybut opisuje tylko jedną encję. Natomiast przy samym odczytywaniu atrybutów nazwa encji może być uwzględniona, np. nr zamówienia, gdzie ZAMÓWIENIE jest nazwą encji, nr jej atrybutem
- Nazwy atrybutów powinny być podane w liczbie pojedynczej – często, gdy nazwy atrybutów podane są w liczbie mnogiej, mamy do czynienia z przypadkiem, gdy powinny to być encje, które mają swoje własne atrybuty.
- Atrybuty powtarzające się powinny być usunięte – w każdej chwili encja może przyjmować tylko jedną wartość dla danego atrybutu. W przypadku, gdy istotnie potrzebna jest większa liczba wartości, powinna zostać utworzona nowa encja, która będzie zawierać te atrybuty w związku wiele do jeden z podstawową encją.

23. Zasady dotyczące atrybutów

PRZYKŁAD. Rozważmy encję ZAMÓWIENIE z następującymi atrybutami

Po zastosowaniu ostatnio wymienionej zasady otrzymujemy następujący diagram



UWAGA. Reguła zastosowana w tym przypadku to w gruncie rzeczy wykorzystanie Drugiej postaci normalnej.

24. Atrybut a encja

- Atrybut powinien być zamodelowany jako encja w diagramie, jeśli jest istotną rzeczą sama w sobie, która posiada własne atrybuty i związki.
- Każda encja musi być jednoznacznie identyfikowana przez atrybut lub kombinację atrybutów (tzw. klucz). Taki zbiór atrybutów powinien być wyznaczony dla każdej encji.
- Należy usunąć te atrybuty, których wartości zależą tylko od części jednoznacznego identyfikatora (sprowadzenie do drugiej postaci normalnej).
- Należy usunąć te atrybuty, których wartości zależą nie tylko od jednoznacznego identyfikatora (sprowadzenie do trzeciej postaci normalnej).
- W niektórych przypadkach atrybut może mieć określoną wartość tylko przez określony przedział czasu, a w niektórych wartości tej w ogóle może nie osiągać. Atrybut taki jest opcjonalny, co oznaczamy przez małe „o” przed nazwą atrybutu.
- W przypadku, gdy atrybut musi zawsze przyjmować określoną wartość, będziemy przed nazwą atrybutu stawiać symbol „*”. Należy jednak zauważyć, iż nie można rozpatrywać wystąpienia encji bez określenia jakiegokolwiek jej wymaganego atrybutu.

25. Zasady dotyczące projektowania – dobór właściwych elementów

W niektórych przypadkach możliwy jest wybór typu elementu projektu, który będzie modelem obiektu ze świata rzeczywistego. Zazwyczaj sytuacja taka sprowadza się do wyboru pomiędzy atrybutem i kombinacją encji i związku. Z reguły łatwiej jest zaimplementować atrybut niż encje lub związek. Jednak modelowanie dowolnych obiektów zawsze jako atrybutu może doprowadzić do pewnych kłopotów.

Przykład. Rozważmy następujący diagram ER.

Czy w powyższym diagramie konieczna jest encja *Wytwórnie*? Być może lepszym rozwiązaniem byłoby dołączenie atrybutów *nazwa*, *adres* i *logo* wytwórni do atrybutów filmu i w ten sposób wyeliminowanie encji *Wytwórnie*. Jednym z powodów wyboru bardziej złożonego elementu, jakim jest encja, był fakt, iż w przeciwnym razie trzeba byłoby powtarzać wielokrotnie atrybuty wytwórni dla każdego filmu. Prowadzi to do redundancji oraz do problemów, jakie mają miejsce w przypadku jej występowania. Jeśli jednak nie ma potrzeby przechowywania adresu i logo wytwórni, to w takim przypadku nazwa wytwórni powinna stać się atrybutem encji *Filmy*, co nie spowoduje powstania redundancji.

26. Reguły rozmieszczania elementów diagramu

Podane poniżej reguły są pomocne w tworzeniu łatwych do odczytania diagramów i umożliwiają ich użytkownikom opracowanie na ich podstawie modelu bazy danych o największej jakości i dokładności.

- Elementy diagramu należy ustawiać w taki sposób, aby ramki encji były wyrównane, a linie związków były prostymi poziomymi lub pionowymi. Należy przy tym zminimalizować liczbę linii, które się przecinają. W przypadku przecinania linii, powinno ono następować pod kątem od 30 do 60 stopni – tak narysowane linie są łatwiejsze do śledzenia i analizowania.
- Powinny być podane tytuł i data oraz dane twórcy każdego diagramu.
- Należy wprowadzać tak tekst, aby był on jednoznaczny. Dlatego należy unikać skrótów i żargonu. Większość tekstu powinna być ustawiona poziomo, w celu ułatwienia jego odczytywania. Dopuszczalne jest wpisanie nazwy obiektu w więcej niż jednym wierszu, aby ułatwić rozmieszczenie elementów diagramu. Jeśli to konieczne, to należy pisać nazwy związków wzdłuż linii, ale zwykle wystarczy umieścić nazwy na końcach linii po przeciwległych jej stronach. Tekst należy konsekwentnie centrować, wyrównywać lewostronnie lub prawostronnie w celu uzyskania jak najlepszej czytelności i przejrzystości wprowadzanych nazw.
- Koniec linii związku o kilku końcówkach (kurza stopka) należy umieszczać z lewej strony linii związku lub u góry linii związku. Dzięki takiemu podejściu wzrasta dokładność modelu poprzez wymuszenie rozpatrywania związków od strony częściej występującej encji w kierunku encji występujących rzadziej. Pomocny jest również fakt, iż przy takim podejściu obiekty, których jest mniej, występują w dolnym prawym rogu diagramu. Są to zazwyczaj encje o dużym znaczeniu, których używa się do definiowania innych obiektów. Zatem analizowanie diagramu w ich kierunku jest bywa pomocne w celu definiowania encji poprzez ich związek z tymi kluczowymi.
- Nie ma większego znaczenia, jaki jest rozmiar i kształt encji. Ramki je obrazujące mogą być dowolnie rozciągnięte, powiększone lub zmniejszone, aby ułatwić narysowanie diagramu.
- Kiedy powyższe konwencje są przestrzegane, to podobne encje zaczynają pojawiać się w sąsiednich miejscach na diagramie.

27. Zasady dotyczące projektowania – unikanie redundancji

Należy unikać wprowadzania powtórzeń i duplikowania tych samych informacji.

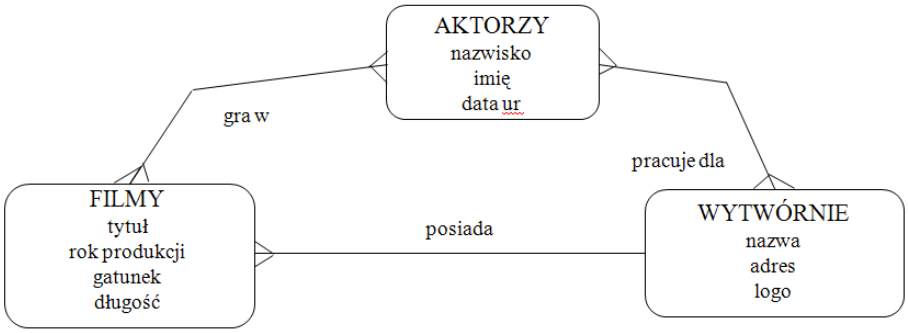
Przykład. Rozważmy przykład encji *Filmy* i *Wytwórnie* pomiędzy, którymi istnieje związek *Posiada*. Encja *Filmy* posiada atrybuty *tytuł*, *rok produkcji*, *gatunek* i *długość*, zaś encja *Wytwórnie* posiada atrybuty *nazwa*, *adres*, *logo*.

Techniczne rzecz ujmując, równie dobrze można by dodać atrybut *nazwa_wytwórni* do zbioru encji *Filmy*. Mimo, iż nie jest to zabronione, to jest to niebezpieczne z kilku powodów:

- Dwa wystąpienia danych informujących o fakcie posiadania filmu przez wytwórnię zajmują więcej miejsca niż pojedyncze wystąpienie.
- Jeśli rozważmy sytuację, w której dochodzi do sprzedaży filmu przez wytwórnię, to może się zdarzyć, iż zostanie zmieniona wytwórnia powiązana przez związek *Posiada*, a nie zostanie zmieniona wartość atrybutu *nazwa_wytwórni* lub też na odwrót, zmieni się wartość atrybutu *nazwa_wytwórni* bez zmiany wartości atrybutu *nazwa*. Oczywiście takich błędów nie powinno się popełniać, tym niemniej w praktyce zdarzają się one często i dlatego duplikowanie informacji jest ewidentnym przyczynkiem, który może spowodować ich pojawienie.

28. Przejście od diagramu związków encji do projektu schematu relacyjnej bazy danych – zamiana encji na relacje

Przykład. Będziemy przekształcać w schemat relacyjnej bazy danych diagram ER postaci:



W pierwszym kroku rozważamy wszystkie encje, które nie są encjami słabymi. Dla każdego takiego typu zbiorów encji należy utworzyć relację o takiej samej nazwie co encja i z takim samym zbiorem atrybutów.

Przykładowo, dla encji *Filmy*:

FILMY			
Tytuł	**Rok produkcji**	**Gatunek**	**Długość**
Gwiezdne wojny	1977	science-fiction	124
Poszukiwacze zaginionej arki	1981	przygodowy	106
Epoka lodowcowa II. Odwilż	2006	kreskówka	104

29. Przejście od diagramu związków encji do projektu schematu relacyjnej bazy danych – zamiana związków na relacje

Na ogół, związki z diagramu ER są konwertowane w schemacie relacyjnej bazy danych na relacje. Taka relacja powstała z danego związku Z ma następujące atrybuty:

- dla każdej encji uczestniczącej w związku Z umieszczamy w schemacie relacji odpowiadającej Z klucze tych encji jako atrybuty tworzonej relacji.
- Jeśli związek ma własne atrybuty, to także te atrybuty są atrybutami tworzonej relacji.
- Jeśli encja uczestniczy w związku wielokrotnie w różnych rolach (każda krawędź symbolizuje inną rolę), to atrybuty kluczy pojawiają się tyle razy, ile jest ról w związku. Jeśli nazwy atrybutów miałyby się powtarzać, należy dokonać ich przemianowania.

Przykład. Rozważmy związek *Posiada* z diagramu prezentowanego na poprzednim slajdzie. W schemacie relacji *Posiada* należy umieścić klucz encji *Filmy*, tj. atrybuty *tytuł* i *rok produkcji* oraz klucz encji *Wytwórnie*, tj. atrybut *nazwa*.

POSIADA		
Tytuł	**Rok produkcji**	**Nazwa**
Gwiezdne wojny	1979	Fox
Poszukiwacze zaginionej arki	1981	Paramount
Epoka lodowcowa II. Odwilż	2006	Fox

30. Przejście od diagramu związków encji do projektu schematu relacyjnej bazy danych – łączenie relacji

W niektórych przypadkach relacje uzyskane z przekształcenia encji i związków nie prowadzą do uzyskania najlepszego schematu relacyjnej bazy danych. Przykładem takiej sytuacji jest przypadek, gdy encja E jest połączona związkiem Z wiele-do-jeden z encją H. Ponadto w schemacie relacji odpowiadającej encji E występują atrybuty encji E, które nie wchodzą w skład klucza, a w relacji powstałej ze związku Z występują atrybuty tworzące klucz encji H oraz własne atrybuty tego związku. Ponieważ związek Z jest typu wiele-do-jeden, a zatem wszystkie wartości atrybutów są jednoznacznie wyznaczone poprzez klucz encji E.

Dlatego można utworzyć połączoną relację o schemacie złożonym z:

- wszystkich atrybutów encji E,
- wszystkich atrybutów encji H,
- wszystkich atrybutów związku Z.

Przykład. Rozważmy związek *Posiada* z diagramu prezentowanego na poprzednim slajdzie, który jest typu wiele-do-jeden pomiędzy encjami *Filmy* i *Wytwórnie*. W przykładzie rozważonym na poprzednim slajdzie został on przekształcony w relację. Relacje dla encji *Filmy* utworzyliśmy we wcześniejszym przykładzie. Okazuje się, iż korzystnym będzie połączyć te dwie relacje, tworząc jedną relację wg przepisu podanego powyżej.

FILMY				
Tytuł	**Rok produkcji**	**Gatunek**	**Długość**	**Nazwa**
Gwiezdne wojny	1977	science-fiction	124	Fox

Poszukiwacze zaginionej arki	1981	przygodowy	106	Paramount
Epoka lodowcowa II. Odwilż	2006	kreskówka	104	Fox

Uwaga. To, w jaki sposób traktować tego typu relacje, zależy od uwarunkowań w modelowanym świecie. Można jednak wskazać zalety pogrupowania wszystkich atrybutów zależnych od klucza encji E w jednej relacji, nawet wtedy, gdy z encji E do innych związków encji prowadzi więcej niż jeden związek typu wiele-do-wielu.

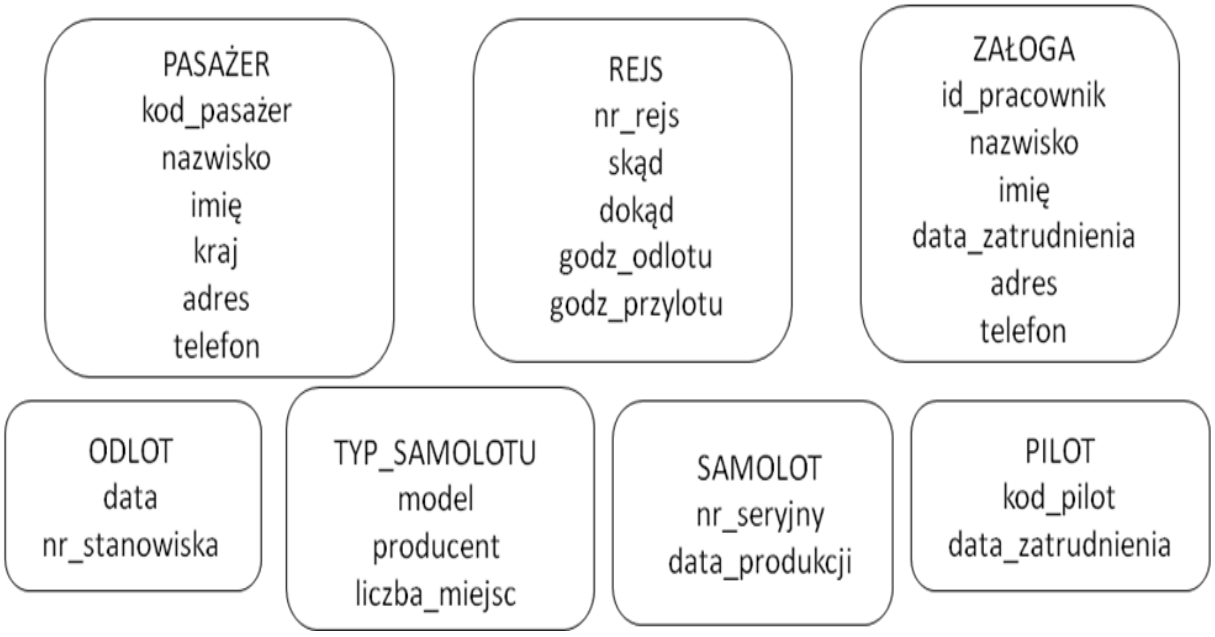
31. Przejście od diagramu związków encji do projektu schematu relacyjnej bazy danych – postępowanie ze zbiorami słabych encji

Jeśli w diagramie związków encji występuje zbiór słabych encji E, to trzeba przestrzegać następujących zasad:

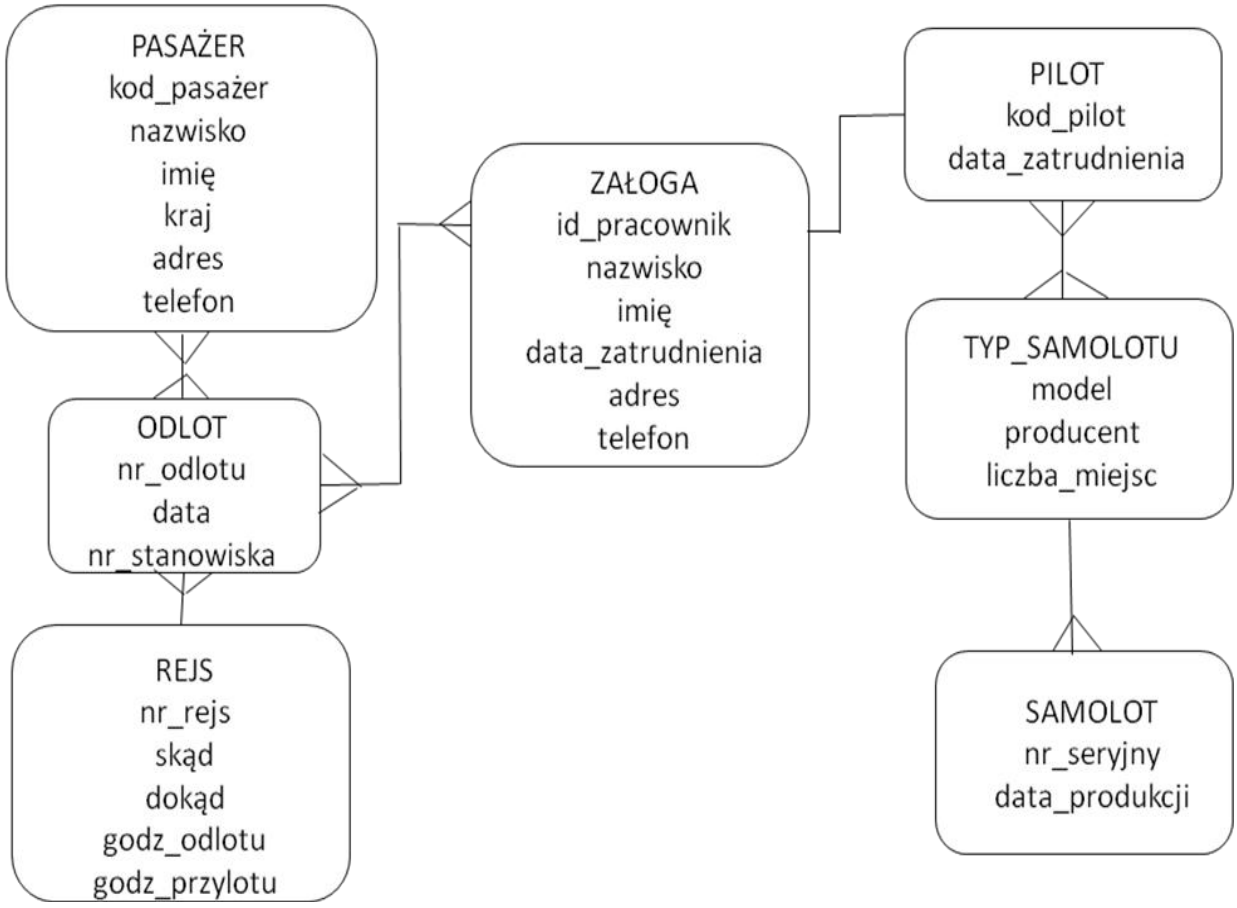
- Relacja odpowiadająca zbiorowi słabych encji E musi zawierać nie tylko atrybuty E, ale także atrybuty kluczy innych zbiorów encji, które wchodzą w skład zbioru E. Te wspomagające zbiory encji są łatwe do rozpoznania, ponieważ uczestniczą wraz z E w związkach wiele-do-jeden, a te związki z kolei są oznaczone podwójnym rombem.
- Każda relacja związku, która obejmuje zbiór słabych encji E, musi korzystać ze wszystkich atrybutów klucza encji E, włączając w to klucze encji E pochodzące z innych zbiorów encji, które uczestniczą w E.
- Z kolei wspierające związki ze zbioru słabych encji E do innego zbioru encji, które ułatwiają zdefiniowanie klucza dla encji E, wcale nie muszą być przekształcane w relację. Atrybuty związku Z typu wiele-do-jeden, o ile je posiada, albo staną się atrybutami relacji odpowiadającej encji E, albo można je połączyć z relacją otrzymaną dla encji E.

32. Przykład diagramu związków encji dla linii lotniczej

1) Najpierw zidentyfikujemy encje mające znaczenie dla rozważanej firmy, tj. linii lotniczej.
   W naszym kręgu zainteresowań znajdują się pasażerowie dokonujący rezerwacji na wybrane loty, samoloty obsługujące te loty oraz personel obsługujący te loty, w tym piloci, którzy kierują samoloty różnych typów, mając w tym zakresie odpowiednie uprawnienia.
   W tym kroku podajemy tylko atrybuty identyfikujące same encje, bez umieszczania atrybutów identyfikujących związki pomiędzy encjami.



2) Następnie wyznaczamy atrybut (zbiór atrybutów) jednoznacznie identyfikujących każdą z powyżej zaprezentowanych encji.
   - ❖ encja PASAŻER – kod_pasażer.
   - ❖ encja REJS – nr_rejs.
   - ❖ encja ZAŁOGA – kod_pracownik.
   - ❖ encja PILOT – kod_pracownik.
   - ❖ encja SAMOLOT – nr_seryjny
   - ❖ encja TYP_SAMOLOT – model
   - ❖ encja ODLOT – nr_odlot.
3) Następnie identyfikujemy związki pomiędzy powyższymi encjami, zapisując je w postaci tekstowej.
   - ❖ PASAŻER *dokonuje rezerwacji* na ODLOT – związek wieloznaczny.
   - ❖ ODLOT *jest* w ramach REJSU – związek jednoznaczny typu jeden-do-wielu (każdy REJS ma jednoznaczny nr rejsu, który jest używany dla wielu odlotów.
   - ❖ ZAŁOGA *przydzielona do* REJSU.
   - ❖ PILOT *wchodzi* w skład ZAŁOGI lotu – związek jednoznaczny typu jeden-do jednego.
   - ❖ PILOT *ma uprawnienia do pilotowania* SAMOLOTU – związek wieloznaczny.
   - ❖ SAMOLOT *jest* określonego TYP SAMOLOTU – związek jednoznaczny typu jeden-do jednego (każdy samolot jest dokładnie jednego typu).
4) Następnie rysujemy diagram uwzględniając wymienione związki.

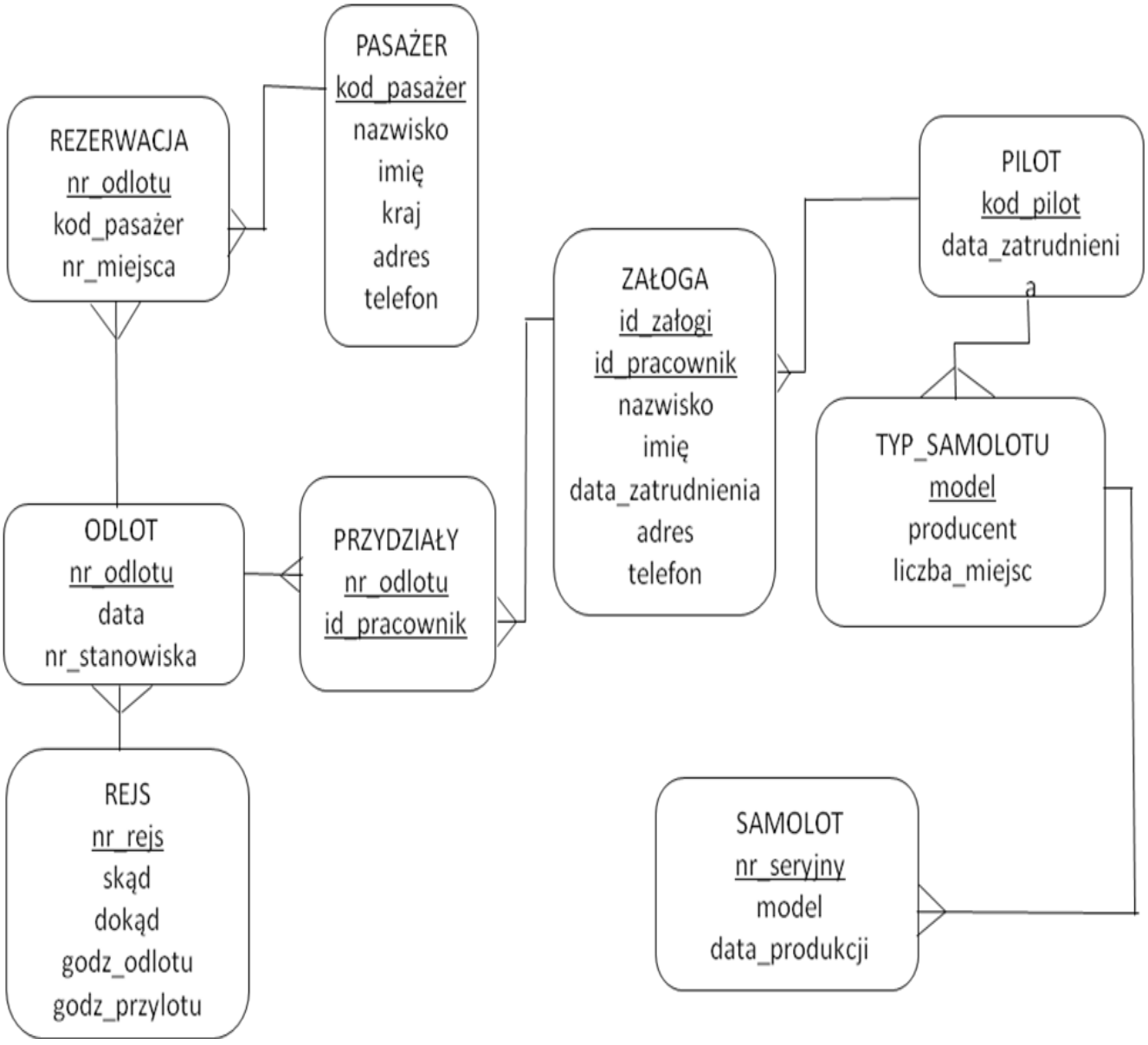


Otrzymaliśmy diagram związków encji, w którym występują związki wieloznaczne:

- ❖ PASAŻER – ODLOT
- ❖ PILOT – TYP_SAMOLOT

- ODLOT – ZAŁOGA

5) Należy teraz wszystkie związki wieloznaczne przekształcić za pomocą dodatkowych encji łącznikowych i par związków jednoznacznych.

- związek *dokonuje rezerwacji na* – zastępujemy nową encją REZERWACJA oraz dwoma związkami:
  - PASAŻER *dokonuje* REZERWACJI,
  - REZERWACJA *dotyczy* ODLOTU,
- związek *przydzielona do* – zastępujemy nową encją PRZYDZIAŁ oraz dwoma związkami:
  - ZAŁOGA *ma* PRZYDZIAŁ,
  - PRZYDZIAŁ *do* ODLOTU,
- związek *ma uprawnienia do pilotowania* – zastępujemy nową encją UPRAWNIENIE oraz dwoma związkami:
  - PILOT *ma* UPRAWNIENIE,
  - UPRAWNIENIE *na określony* TYP_SAMOLOTU.

6) Otrzymujemy wynikowy schemat