

Plan

- Model relacyjny
- Algebra relacji
- Normalizacja
- Mapowanie modelu koncepcyjnego ER na relacyjny
- Mapowanie modelu koncepcyjnego EER na model relacyjny

Model relacyjny

- Podstawowe pojęcia
- Formalne definicje
- Rodzaje kluczy
- Organiczenia relacyjne (więzy integralności)

Podstawowe pojęcia

- Model relacyjny po raz pierwszy sformalizowany przez Edgara F. Codd'a w 1970 roku
 - E.F. Codd, *A relational model of data for large shared data banks*. Communications of the ACM (1970)
- Formalny model danych z solidnymi podstawami matematycznymi, oparty o teorię zbiorów i logikę predykatów pierwszego rzędu
- Brak reprezentacji graficznej
- Powszechnie stosowany do budowy zarówno logicznych, jak i wewnętrznych modeli danych
Microsoft SQL Server, IBM DB2, Oracle

Podstawowe pojęcia

- Baza danych jest reprezentowana jako zbiór relacji
- Relację definiuje się jako zbiór krotek, z których każda reprezentuje podobną encję świata rzeczywistego
- Krotka to uporządkowana lista wartości atrybutów, z których każda opisuje aspekt encji

Podstawowe pojęcia

SUPPLIER

SUPNR	SUPNAME	SUPADDRESS	SUPCITY	SUPSTATUS
21	Deliwines	240, Avenue of the Americas	New York	20
32	Best Wines	660, Market Street	San Francisco	90
37	Ad Fundum	82, Wacker Drive	Chicago	95
52	Spirits & co.	928, Strip	Las Vegas	NULL
68	The Wine Depot	132, Montgomery Street	San Francisco	10
69	Vinos del Mundo	4, Collins Avenue	Miami	92
...				

Podstawowe pojęcia

Model ER	Model relacyjny
Typ encji	Relacja
Encja	Krotka
Typ atrybutu	Nazwa kolumny
Atrybut	Komórka

Postawowe pojęcia

Student (Studentnr, Name, HomePhone, Address)

Professor (SSN, Name, HomePhone, OfficePhone, E-mail)

Course (CourseNo, CourseName)

Formalne definicje

- Dziedzina określa zakres dopuszczalnych wartości dla typu atrybutu
 - np.: dziedzina dla płci, dziedzina dla czasu
- Każdy typ atrybutu jest zdefiniowany przy użyciu odpowiedniej dziedziny
- Dziedzina może być używana wielokrotnie w relacji

BillOfMaterial

MAJORPRODNR	MINORPRODNR	QUANTITY
5	10	2
10	15	30

Formalne definicje

- *Relacja* $R(A_1, A_2, A_3, \dots A_n)$ może być formalnie zdefiniowana jako zbiór m krotek $r = \{t_1, t_2, t_3, \dots t_m\}$ gdzie każda *krotka* t jest uporządkowaną listą n wartości $t = \langle v_1, v_2, v_3, \dots v_n \rangle$ odpowiadającą określonej encji
 - każda wartość v_i jest elementem odpowiedniej dziedziny, $\text{dom}(A_i)$, lub jest specjalną wartością NULL
 - Wartość NULL oznacza, że brakuje wartości, jest ona nieistotna lub nie dotyczy

Formalne definicije

Student(100, Michael Johnson, 123 456
789, 532 Seventh Avenue)

Professor(50, Bart Baesens, NULL, 876
543 210, Bart.Baesens@kuleuven.be)

Course(10, Principles of Database
Management)

Formalne definicje

- Relacja zasadniczo reprezentuje zbiór (bez porządkowania + bez duplikatów)
- Ograniczenie dziedziny stanowi, że wartość każdego typu atrybutu A musi być niepodzielną i pojedynczą wartością z dziedziny $\text{dom}(A)$
- np.: `COURSE(coursenr, coursename, study points)`

`(10, Principles of Database Management, 6)`

`(10, {Principles of Database Management, Database Modeling}, 6) → BŁĄD`

Formalne definicje

- Relacja R stopnia n na dziedzinach $\text{dom}(A_1)$, $\text{dom}(A_2)$, $\text{dom}(A_3)$, ..., $\text{dom}(A_n)$ może być również alternatywnie zdefiniowana jako podzbiór iloczynu kartezyjskiego dziedzin, które definiują każdy z typów atrybutów

Domain Product ID
001
002
003
...

X

Domain Product Color
Blue
Red
Black

X

Domain Product Category
A
B
C

ProductID	Product Color	Product Category
001	Blue	A
001	Blue	B
001	Blue	C
001	Red	A
001	Red	B
001	Red	C
...		

Rodzaje kluczy

- Nadklucze (superklucze) i klucze
- Klucze kandydujące, klucze główne, klucze alternatywne
- Klucze obce

Nadklucze i klucze

- Nadklucz jest zdefiniowany jako podzbiór typów atrybutów relacji R z właściwością, że żadne dwie krotki w jakimkolwiek stanie relacji nie powinny mieć tej samej kombinacji wartości dla tych typów atrybutów
- Nadklucz określa ograniczenie unikalności
- Nadklucz może mieć nadmiarowe typy atrybutów
- np.: (Studentnr, Name, HomePhone)

Nadklucze i klucze

- Klucz K schematu relacji R jest nadkluczem R z dodatkową właściwością, że usunięcie dowolnego typu atrybutu z K pozostawia zestaw typów atrybutów, które nie są kluczem R
- Klucz nie ma żadnych nadmiarowych typów atrybutów (minimalny nadklucz)
- np.: Studentnr
- Ograniczenie klucza stanowi, że każda relacja musi mieć co najmniej 1 klucz, który pozwala jednoznacznie zidentyfikować jej krotki

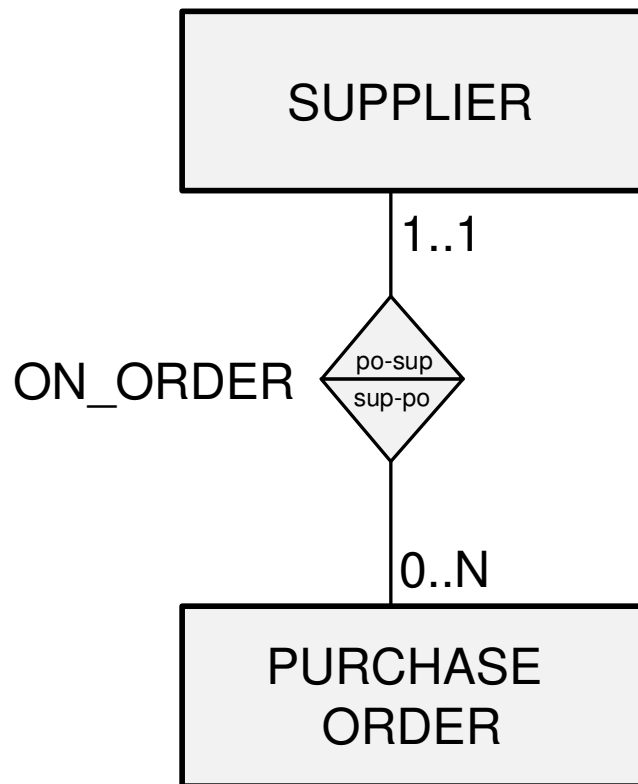
Klucze kandydujące, klucze główne i klucze alternatywne

- Relacja może mieć więcej niż jeden klucz (klucze kandydujące)
 - PRODUCT: product numer i product name
- Klucz główny (primary key) służy do identyfikowania krotek w relacji, ustalania połączeń z innymi relacjami oraz do celów przechowywania
 - Ograniczenie integralności encji: typy atrybutów, które tworzą klucz główny, powinny zawsze spełniać ograniczenie NOT NULL
- Inne klucze kandydujące są wtedy nazywane kluczami alternatywnymi

Klucze obce

- Zbiór typów atrybutów FK w relacji R_1 jest kluczem obcym R_1 jeżeli spełnione są dwa warunki (ograniczenie integralności referencyjnej - referential integrity constraint)
 - typy atrybutów w FK mają te same dziedziny, co typy atrybutów klucza głównego PK relacji R_2
 - wartość FK w krotce t_1 w bieżącym stanie r_1 albo występuje jako wartość PK dla pewnej krotki t_2 w bieżącym stanie r_2 albo jest NULL

Klucze obce



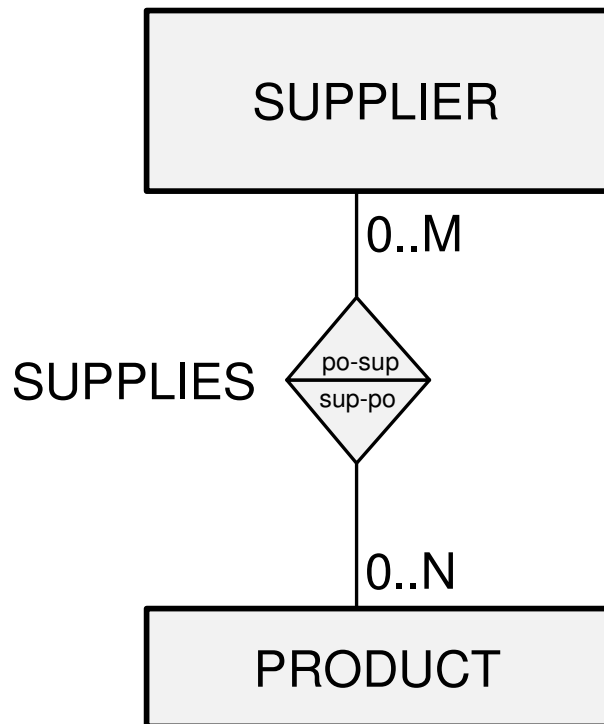
SUPPLIER

SUPNR	SUPNAME	SUPADDRESS	SUPCITY	SUPSTATUS
...				
37	Ad Fundum	82, Wacker Drive	Chicago	95
94	The Wine Crate	330, McKinney Avenue	Dallas	75
...				

PURCHASE_ORDER

PONR	PODATE	<i>SUPNR</i>
1511	2015-03-24	37
1512	2015-04-10	94
...		

Klucze obce



SUPPLIER

SUPNR	SUPNAME	SUPADDRESS	SUPCITY	SUPSTATUS
21	Deliwines	240, Avenue of the Americas	New York	20
32	Best Wines	660, Market Street	San Francisco	90
...				

PRODUCT

PRODNR	PRODNAME	PRODTYPE	AVAILABLE_QUANTITY
0119	Chateau Miraval, Cotes de Provence Rose, 2015	rose	126
0154	Chateau Haut Brion, 2008	red	111
...		red	5

Klucze obce

SUPPLIES

<u>SUPNR</u>	<u>PRODNR</u>	PURCHASE_PRICE	DELIV_PERIOD
...			
68	0327	56.99	4
...			
21	0289	17.99	1
21	0327	56.00	6
21	0347	16.00	2
...			
69	0347	18.00	4
84	0347	18.00	4
...			

Ograniczenia relacyjne

Ograniczenie dziedziny	Wartość każdego typu atrybutu A musi być niepodzielną i pojedynczą wartością z dziedziny $dom(A)$.
Ograniczenie klucza	Każda relacja ma klucz, który pozwala jednoznacznie zidentyfikować jej krotki.
Ograniczenie integralności encji	Typy atrybutów, które tworzą klucz główny, powinny zawsze spełniać ograniczenie NOT NULL.
Ograniczenie integralności referencyjnej	Klucz obcy FK ma tę samą dziedzinę, co typ(y) atrybutu PK klucza głównego, do którego się odwołuje i występuje jako wartość PK lub NULL.

Przykład modelu relacyjnego

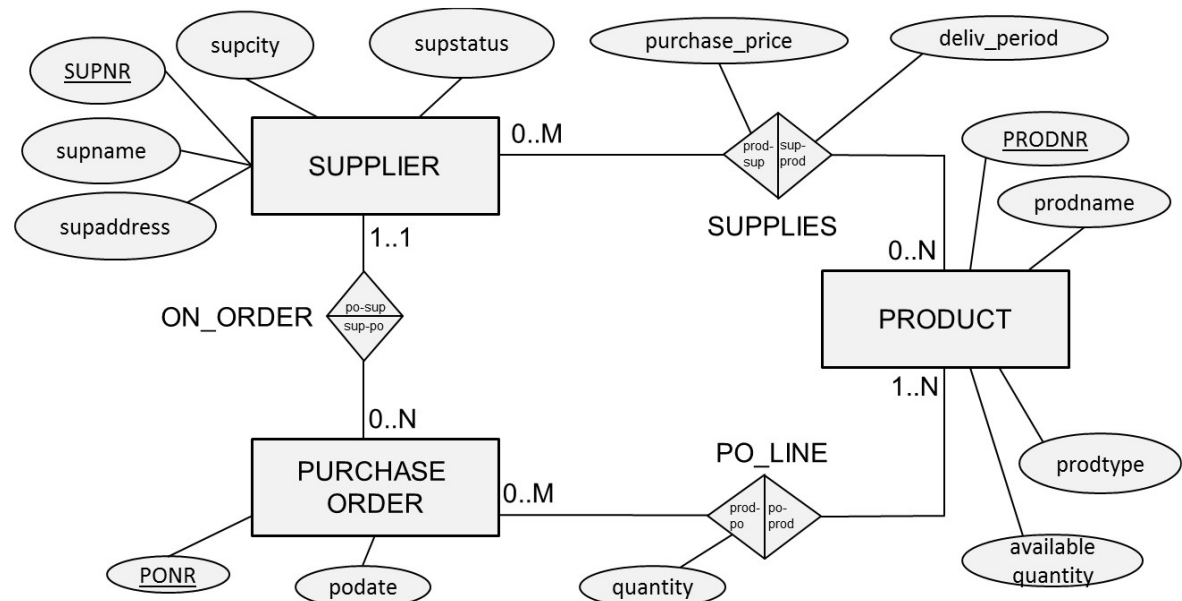
SUPPLIER(SUPNR, SUPNAME, SUPADDRESS, SUPCITY, SUPSTATUS)

PRODUCT(PRODNR, PRODNAME, PRODTYPE, AVAILABLE QUANTITY)

SUPPLIES(SUPNR, PRODNR, PURCHASE_PRICE, DELIV_PERIOD)

PURCHASE_ORDER(PONR, PODATE, *SUPNR*)

PO_LINE(PONR, PRODNR, QUANTITY)



Przykład modelu relacyjnego

Supplier

SUPNR	SUPNAME	SUPADDRESS	SUPCITY	SUPSTATUS
21	Deliwines	240, Avenue of the Americas	New York	20
32	Best Wines	660, Market Street	San Francisco	90
...				

Product

PRODNR	PRODNAME	PRODTYPE	AVAILABLE_QUANTITY
0119	Chateau Miraval, Cotes de Provence Rose, 2015	rose	126
0384	Dominio de Pingus, Ribera del Duero, Tempranillo, 2006	red	38
...			

Supplies

SUPNR	PRODNR	PURCHASE_PRICE	DELIV_PERIOD
21	0119	15.99	1
21	0384	55.00	2
...			

Purchase_Order

PONR	PODATE	SUPNR
1511	2015-03-24	37
1512	2015-04-10	94
...		

PO_Line

PONR	PRODNR	QUANTITY
1511	0212	2
1511	0345	4
...		