

РЕЛАЦИОННИ БАЗИ ОТ ДАННИ

Релационен модел на данните

- ▶ Създаден през 1970 от Dr. E.F. Codd, специалист на IBM
- ▶ Развит от Chris Date и Hugh Darwen
- ▶ 1985 - Codd публикува списък от 13 правила, познати като "Codd's 12 Rules"
- ▶ 1995 – публикуван е 3-тия манифест, в който Chris Date и Hugh Darwen показват как релационният модел придобива обектно-ориентирани възможности

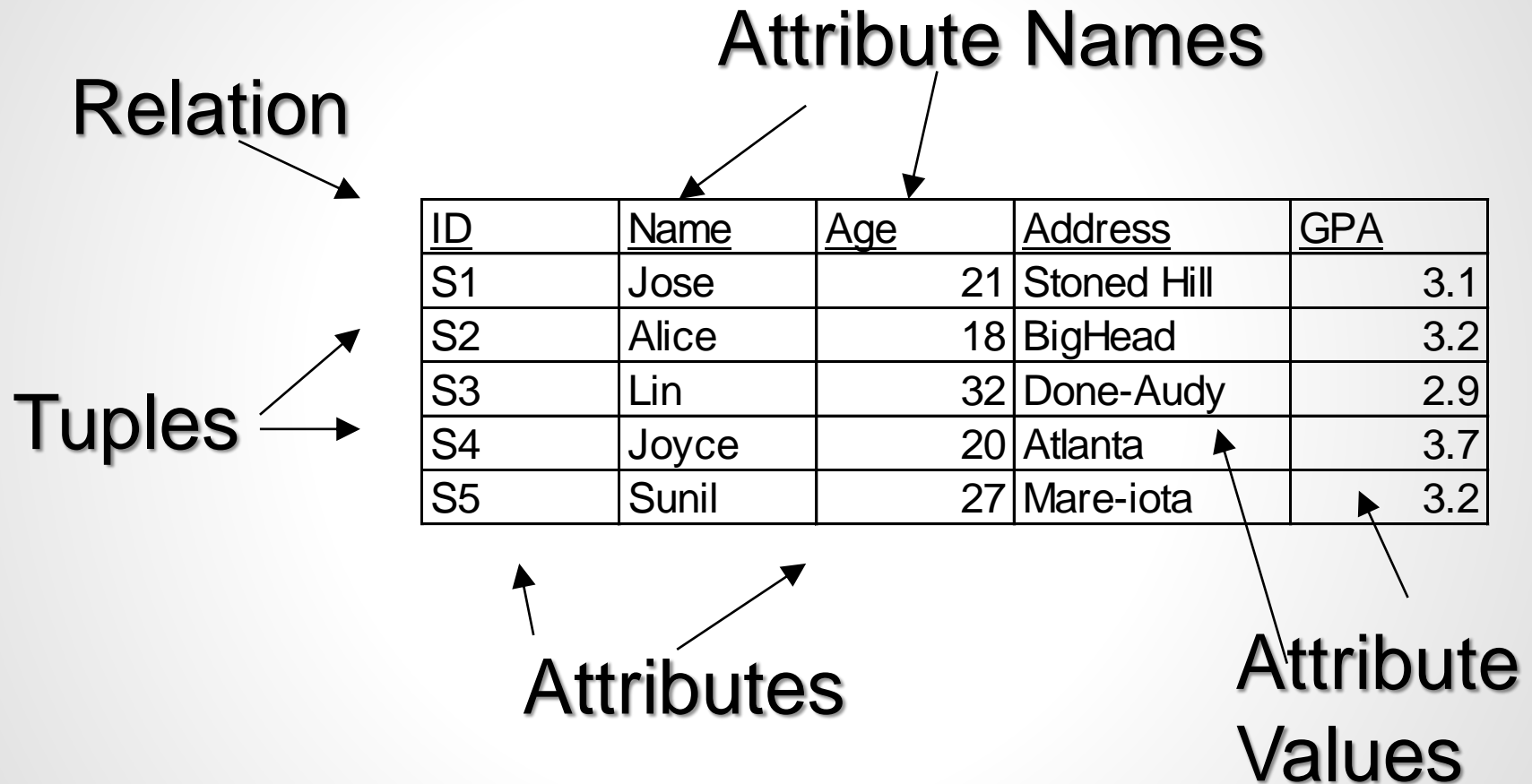
Релационна теория на Е.Код

- ▶ Фундаментално предположение в релационната теория
 - всички данни се представят като математически n -арни релации, които са подмножество от картезианско умножение на n домейна.
- ▶ **Релация $A = (a_1, a_2, \dots, a_n)$**
- ▶ Оперирането с данните се осъществява по правилата на релационните изчисления или релационната алгебра.

Определение на релация

- ▶ Релацията (relation) е съвкупност от кортежи (tuples) (d_1, d_2, \dots, d_j) , където всеки елемент d_n е член на множество D_n , наречено **домейн** (data domain).
- ▶ Всеки отделен домейн, използван в дефиницията на релацията, се нарича **атрибут** (attribute) и всеки атрибут е именуван.
- ▶ Наборът от атрибути (имената на атрибутите с техните асоциирани домейни) се нарича **релационна схема** (relation schema или relation scheme).

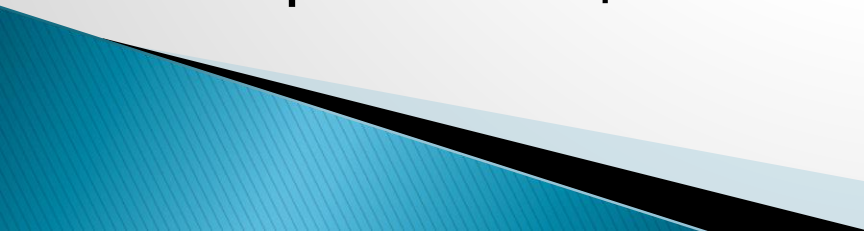
Релация (релационна таблица)




Формални и неформални термини в релационния модел

Формални	Неформални	Други
Релация (Relation)	Таблица (Table)	Таблица (Table)
Кортеж (Tuple)	Ред (Row)	Запис (Record)
Атрибут (Attribute)	Колона (Column)	Поле (Field)

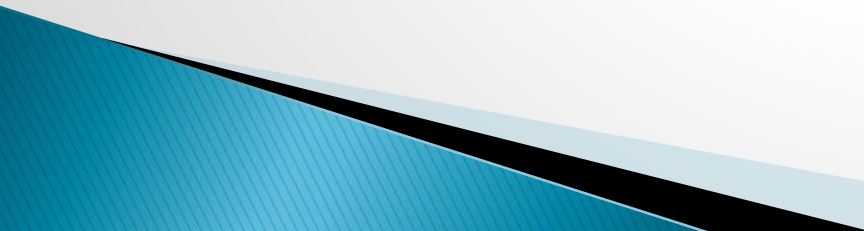
Основни концепции в релационния модел

- ▶ Структури от данни и терминология
 - ▶ Нотации
 - ▶ Свойства на релационните таблици
 - ▶ Връзки и ключове
 - ▶ Интегритет на данните
 - ▶ Релационни операции с данните
 - ▶ Нормализация
- 

Структури от данни и терминология

- ▶ Релационният модел представя данните в двумерни таблици.
 - ▶ Всяка таблица представя някакви реално съществуващи хора, места, неща или събития, за които е събрана информация.
 - ▶ Релационната база от данни е съвкупност (колекция) от двумерни таблици, които са свързани помежду си.
- 

Свойства на релацията

- ▶ Атомарни стойности
 - ▶ Стойностите в колоната са от един и тип
 - ▶ Всеки ред е уникален
 - ▶ Последователността на колоните не е от значение
 - ▶ Последователността на редовете не е от значение.
 - ▶ Всяка колона има уникално име
- 

AUTHOR

au_id	au_lname	au_fname	address	city	state
172-32-1176	White	Johnson	10932 Bigge Rd.	Menlo Park	CA
213-46-8915	Green	Marjorie	309 63rd St. #411	Oakland	CA
238-95-7766	Carson	Cheryl	589 Darwin Ln.	Berkeley	CA
267-41-2394	O'Leary	Michael	22 Cleveland Av. #14	San Jose	CA
274-80-9391	Straight	Dean	5420 College Av.	Oakland	CA
341-22-1782	Smith	Meander	10 Mississippi Dr.	Lawrence	KS
409-56-7008	Bennet	Abraham	6223 Bateman St.	Berkeley	CA
427-17-2319	Dull	Ann	3410 Blonde St.	Palo Alto	CA
472-27-2349	Gringlesby	Burt	PO Box 792	Covelo	CA
486-29-1786	Looksley	Charlene	18 Broadway Av.	San Francisco	CA

TITLE

title_id	title	type	price	pub_id
BU1032	The Busy Executive's Database Guide	business	19.99	1389
BU1111	Cooking with Computers	business	11.95	1389
BU2075	You Can Combat Computer Stress!	business	2.99	736
BU7832	Straight Talk About Computers	business	19.99	1389
MC2222	Silicon Valley Gastronomic Treats	mod_cook	19.99	877
MC3021	The Gourmet Microwave	mod_cook	2.99	877
MC3026	The Psychology of Computer Cooking	UNDECIDED		877
PC1035	But Is It User Friendly?	popular_comp	22.95	1389
PC8888	Secrets of Silicon Valley	popular_comp	20	1389
PC9999	Net Etiquette	popular_comp		1389
PS2091	Is Anger the Enemy?	psychology	10.95	736

PUBLISHER

pub_id	pub_name	city
736	New Moon Books	Boston
877	Binnet & Hardley	Washington
1389	Algodata Infosystems	Berkeley
1622	Five Lakes Publishing	Chicago
1756	Ramona Publishers	Dallas
9901	GGG&G	München
9952	Scootney Books	New York
9999	Lucerne Publishing	Paris

AUTHOR TITLE

au_id	title_id
172-32-1176	PS3333
213-46-8915	BU1032
213-46-8915	BU2075
238-95-7766	PC1035
267-41-2394	BU1111
267-41-2394	TC7777
274-80-9391	BU7832
409-56-7008	BU1032
427-17-2319	PC8888
472-27-2349	TC7777

Схема на релацията

► $R (R_1, R_2, R_3, \dots, R_n)$

AUTHOR	(au_id , au_lname, au_fname, address, city, state, zip)
TITLE	(title_id , title, type, price, pub_id)
PUBLISHER	(pub_id , pub_name, city)
AUTHOR_TITLE	(au_id , title_id)

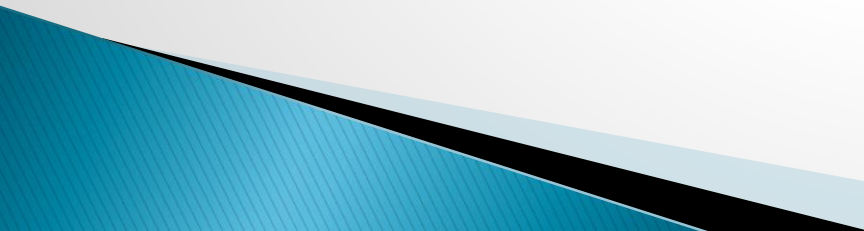
Връзки (relationships) и ключове (keys)

- ▶ Връзки (relationships)
- ▶ Ключове (keys)
 - Първичен ключ (Primary Key – PK)
 - Външен ключ (Foreign key - FK)
 - Възможен ключ - колона или група от колони, които уникално идентифицират редовете в релацията

Първичен ключ (Primary Key - PK)





- ❑ Колона или група от колони, които са избрани да идентифицират редовете в релацията
- ❑ Изисквания към РК
 - Уникалност – не може да приема дублиращи се стойности
 - Is No Null - не може да приема неопределени стойности
 - Минималност – нито един от атрибутите, влизащи в състава на РК не може да се премахне без това да наруши свойството уникалност

Роля на първичния ключ

- ▶ Уникално определя редовете в релацията
 - ▶ По него се търсят данните в релацията
 - ▶ Редовете в релацията се подреждат във възходящ ред по неговите значения при показване на съдържанието на релацията
 - ▶ Повечето СУБД по подразбиране създават индекс по първичния ключ (обикновено, първичен индекс)
 - ▶ Използва се за връзка с други релации
- 

Индекс (Index)

- Индексът е нареден набор от указатели към данните в таблицата
- Повишава производителността – достъпът до данните е по-бърз
- Първични и вторични индекси

E-No	Pointer
179	
342	
719	
857	

Employee

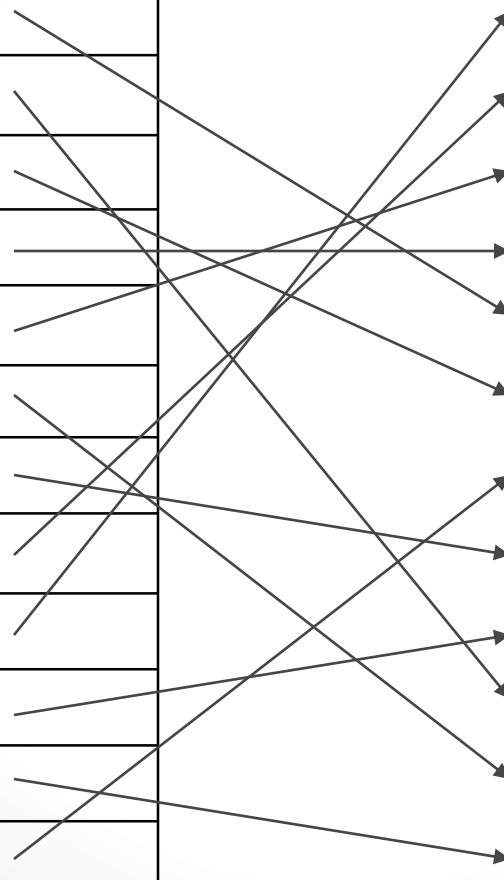
E-No	E-Name	D-No
179	Silva	7
857	Perera	4
342	Dias	7
719	De Silva	5

Index: Employee Name

E-Name	Pointer
Alwis	
Bandara	
Costa	
De Silva	
Dias	
Opatha	
Peiris	
Perera	
Silva	
Vaas	
Wickrama	
Zoysa	

Employee

E-No	E-Name	D-No
179	Silva	7
857	Perera	4
342	Dias	7
719	De Silva	5
587	Alwis	4
432	Costa	6
197	Zoysa	2
875	Peiris	4
324	Vaas	7
917	Bandara	3
785	Opatha	2
234	Wickrama	1



Връзки между релациите

- ▶ На основата на общи съпоставими атрибути (колони).
- ▶ Пример:

Employee(Emp_No, Emp_Name, *Department*)

Department(Dept_No, Dept_Name, M_No)

- ▶ Атрибут в една релация, който съответства на първичния ключ в друга релация наричаме **ВЪНШЕН КЛЮЧ** (Foreign Key - FK).

За потребителите данните се представят като таблици!

- Външният ключ е колона или група от колони в една таблица, които служат като първичен ключ в друга таблица.

Employee

E-No	E-Name	D-No
179	Silva	7
857	Perera	4
342	Dias	7

Primary Key

Foreign Key

Department

D-No	D-Name	M-No
4	Finance	857
7	Sales	179

Primary Key

Employee - таблица родител, Department - свързана (подчинена) таблица

Връзки между релациите Employee

E-No	E-Name	D-No
179	Silva	7
857	Perera	4
342	Dias	7

Primary Key

Foreign Key

Редовете в една или повече
таблици се асоциират една с
друга чрез стойностите в
колоните (не чрез указатели).

Department

D-No	D-Name	M-No
4	Finance	857
7	Sales	179

Primary Key

Foreign Key

Salary

E-No	Eff-Date	Amt
179	1/1/98	8000
857	3/7/94	9000
179	1/6/97	7000
342	28/1/97	7500

Foreign Key

Primary Key

Рекурсивен външен ключ: Външен ключ в релация, който съответства на първичния ключ в същата релация.

Пример:

E-No	E-Name	D-No	Manager-No
179	Silva	7	
857	Perera	4	179
342	Dias	7	179

Primary Key

Foreign Key
рекурсивен

Интегритет (integrity) на данните

- ▶ *Дадена съвкупност от данни притежава свойството **интегритет (цялост)**, ако данните са логически съвместими, непротиворечиви и изпълняват определени ограничения.*
- ▶ Основни типове интегритет на данните
 - Entity Integrity
 - Domain Integrity
 - Referential Integrity

Интегритет на данните

- ▶ **Интегритет по същност (Entity Integrity)**
 - Релацията да притежава първичен ключ
- ▶ **Домейн интегритет (Domain Integrity)**
 - Дефинира се за отделни атрибути на релацията
 - Специфицира набор от стойности, които са валидни за даден атрибут (дадена колона) и дали се разрешава атрибутът да приема неопределени стойности

Пример: В таблица „Изпити“ за колоната „Оценка“ множеството допустими стойности са целите числа от 2 до 6 и не може да не приема стойност.

Интегритет на връзките (Referential Integrity)

- ▶ Дефинира се за релации, между които съществува връзка на основата на общи атрибути (първичен и външен ключове)
- ▶ Изисква външният ключ в дадена релация да приема
 - стойности само измежду съществуващите стойности на първичния ключ в другата релация
 - или неопределена стойност (NULL).

Интегритетът на връзките налага определени правила при

- ▶ Добавянето на нови редове в породената таблица;
- ▶ Промяна стойността на първичния ключ в таблицата родител;
- ▶ Изтриване на редове от таблицата родител.

Правила при добавяне на нови редове в свързаната таблица(Insert Rules)

- ▶ **Dependent** – добавя се нов ред в свързаната таблица само, ако стойността на FK съществува като стойност на първичния ключ в таблицата родител
- ▶ **Nullify**
- ▶ Automatic
- ▶ Default
- ▶ Customized
- ▶ No Effect – не се контролира интегритета

Правила при изтриване на редове от таблицата родител (Delete Rules)

- ▶ Определят какви промени се правят в свързаната (породената) таблица
 - **Restrict** – не се разрешава изтриване на ред от таблицата родител, ако има свързани редове в свързаната таблица.
 - **Cascade**
 - Nullify
 - Default
 - Customized
 - No Effect

Правила при обновяване на РК в таблицата родител (Update Rules)

- ▶ Определят дали може да се променят съществуващите стойности на РК в таблицата родител и влиянието върху FK в свързаната таблица
 - **Restrict** – не се разрешава промяна стойност на РК в ред от таблицата родител, ако има съответни редове в свързаната таблица.
 - **Cascade (каскадно обновяване)** – при промяна стойността на РК в таблицата родител автоматично се променят стойностите на външния ключ в съответните редове на породената таблица.
 - Nullify
 - Default
 - Customized
 - No Effect

Релационни правила на д-р Е. Код

- 0. Принцип на създаване
 - 1. Информация
 - 2. Гарантиран достъп
 - 3. Поддържане на неопределени стойности (Null Value)
 - 4. Активен речник (каталог) на данните
 - 5. Пълен (изчерпателен) език за описание и манипулиране с БД

Релационни правила на д-р Е. Код

6. Актуализиране (обновяване) на данните чрез изгледи
 7. Актуализация на ниво множество (набори от данни)
 8. Физическа независимост на данните
 9. Логическа независимост на данните
 10. Независимост на интегритета на данните
 11. Независимост на разпространението
 12. Устойчивост
- 