Relational Data Model

CE3101 - Bases de Datos

Disclaimer / Descargo de Responsabilidad

Esta presentación corresponde a una guía usada por el profesor durante las clases. La misma ha sido modificada para ser utilizado en el modelo de cursos asistidos por tecnología. No es una versión final, por lo que la misma podría requerir todavía hacer algunos ajustes. Para aspectos de evaluación esta presentación es solo una guía, por lo que el estudiante debe profundizar con el material de lectura asignado y lo discutido en clases para aspectos de evaluación.

This presentation corresponds to a guide material used by the professor during classes. It has been modified to be used in the model of technology-assisted courses. It is not a final version, so it may still require some adjustments. For evaluation aspects, this presentation is only a guide, so the student should delve with the assigned reading material and what has been discussed in class.

Pregunta inicial

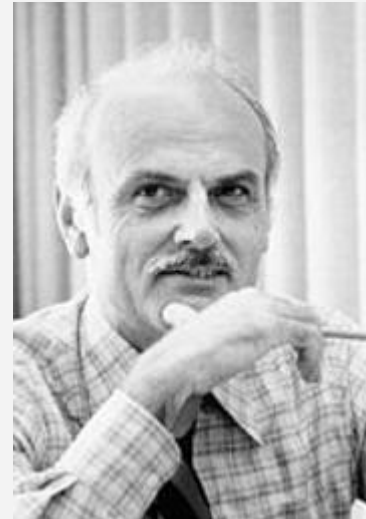Los perros ladran, los gatos maúllan, los pájaros pían.

Que hacen las jirafas?

The Relational Data Model

→ The relational data model was first introduced by Ted Codd of IBM Research in 1970.
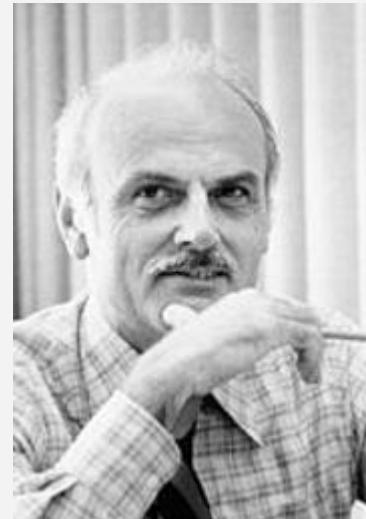
The Relational Data Model

➔ The relational data model was first introduced by Ted Codd of IBM Research in 1970.

The Relational Data Model

➔ The relational data model was first introduced by Ted Codd of IBM Research in 1970.

➔ The model uses the concept of a mathematical relation.

➔ The first commercial implementation of the relational model became available in the early 1980.

Relational Model Concepts

CONCEPTS	
➜ The relational model represents the database as a collection of **relations**.	
➜ Informally, each relation resembles a table of values or, to some extent, a flat file of records.	
➜ A row represent a fact that typically corresponds to a real-world entity or relationship.	

Relational Model Concepts

CONCEPTS	
➜ The relational model represents the database as a collection of **relations**.	
➜ Informally, each relation resembles a table of values or, to some extent, a flat file of records.	
➜ A row represent a fact that typically corresponds to a	

The table name and column names are used to help to interpret the meaning of the values in each row.

Relational Model Concepts

CONCEPTS

➜ Represents the database as a collection of **relations**.

➜ Each relation resembles a table of values or flat file of records.

➜ A row represent a fact that typically corresponds to a real-world entity or relationship.

TERMINOLOGY

In the formal relational model terminology:

➜ A **row** is called a *tuple*.

➜ A **column header** is called an *attribute*.

➜ A **table** is called a *relation*.

Domains, Attributes, Tuples and Relations

DOMAIN

➔ A domain D is a **set of atomic values**.
  ◆ Names: The set of character strings that represents names of personas.
  ◆ Employee ages: Possible ages of employee in a company; each must be an integer value between 15 and 80.
 	
➔ A **data type** or **format** is also specified for each domain.

Domains, Attributes, Tuples and Relations

DOMAIN

→ A domain D is a **set of atomic values**.
  ◆ Names: The set of character strings that represents names of personas.
  ◆ Employee ages: Possible ages of employee in a company; each must be an integer value between 15 and 80.

→ A **data type** or **format** is also specified for each domain.

RELATION

→ A relation schema **R**, denoted by **R(A1, A2,…, An)**, is made up of a relation name **R** and a list of attributes **A1, A2,…, An**. **E**ach attribute **Ai** is the name of a role played by some domain **D** in the relation schema **R**.
→ The degree of a relation is the number of attributes n of its relation schema.
→ Using the data type of each attribute.

STUDENT(Name: string, Ssn: string, Home_phone: string, Address: string, Office_phone: string, Age: integer, Gpa: real)

Domains, Attributes, Tuples and Relations

Relation Name

Attributes

STUDENT

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	(817)373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	(817)375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	(817)749-1253	25	3.53
Rohan Panchal	489-22-1100	(817)376-9821	265 Lark Lane	(817)749-6492	28	3.93
Barbara Benson	533-69-1238	(817)839-8461	7384 Fontana Lane	NULL	19	3.25

Tuples

Characteristics of Relations

➔ **Ordering of tuples in a relation.** Elements of a set have no order among them.

➔ **Ordering of values within a tuple.** An n-tuple is an ordered list of n values.

Characteristics of Relations
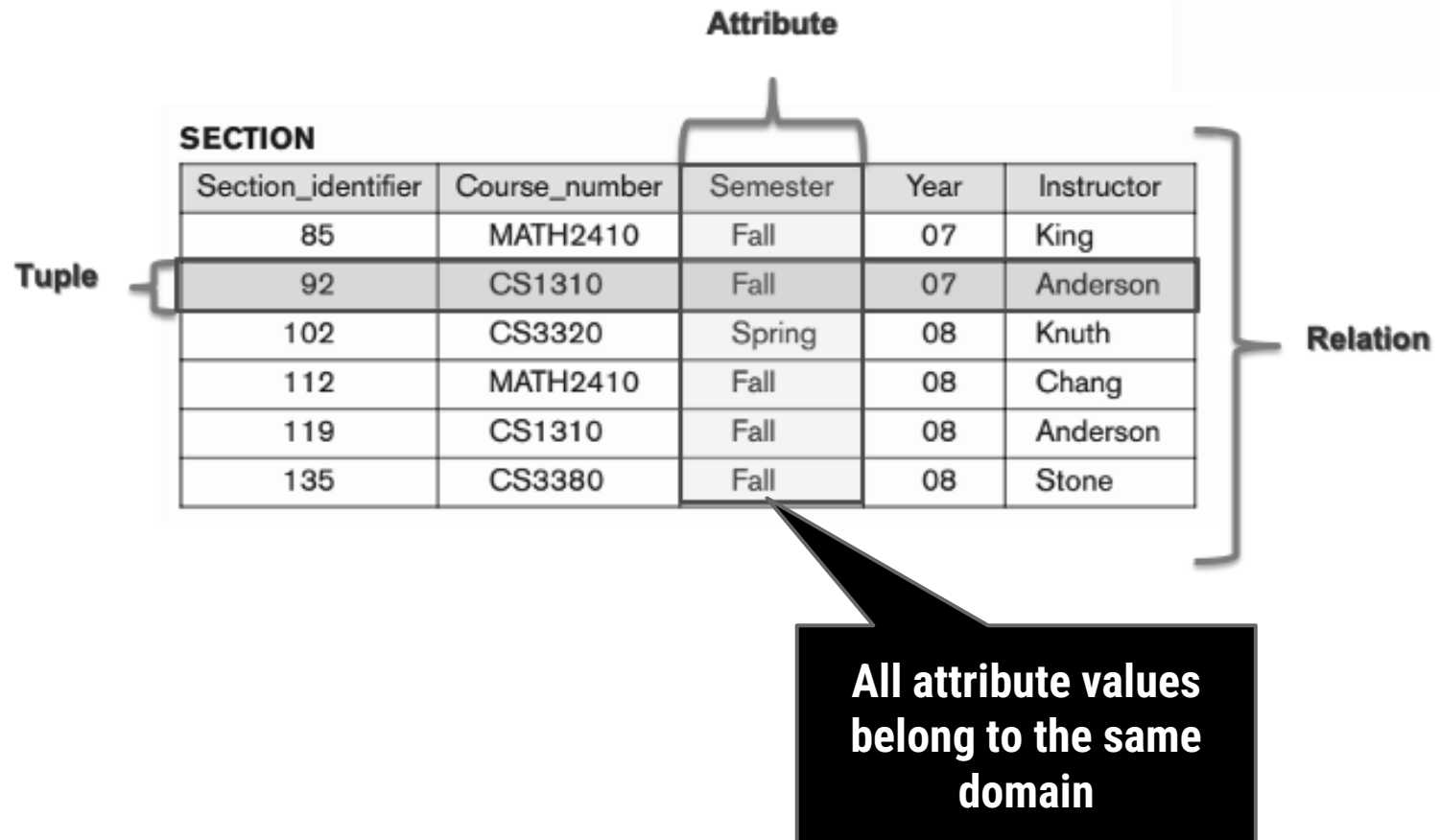
- → **Ordering of tuples in a relation.** Elements of a set have no order among them.

- → **Ordering of values within a tuple.** An n-tuple is an ordered list of n values.

- → **Values and NULL is the tuples.** Each value in a tuple is an **atomic value**.

- → **Interpretation (meaning) of a relation.** The relation schema can be interpreted as a declaration or a type of assertion.

Essential Concepts

Attribute

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	07	King
92	CS1310	Fall	07	Anderson
102	CS3320	Spring	08	Knuth
112	MATH2410	Fall	08	Chang
119	CS1310	Fall	08	Anderson
135	CS3380	Fall	08	Stone

Tuple

Relation

All attribute values belong to the same domain

Relational Model Constraints

➜ It is a feature of a database using the relational model.

➜ Three categories:

  ◆ Inherent model-based constraints or implicit constraints

  ◆ Schema-based constraints or explicit constraints

  ◆ Application-based or semantic constraints or business rules

Inherent model-based constraints or implicit constraints

DEFINITION

→ Restrict the values that can be stored considering the model definition.

Inherent model-based constraints or implicit constraints

DEFINITION

➜ Restrict the values that can be stored considering the model definition.

EXAMPLE

➜ A relation cannot have duplicate tuples is an inherent constraint.

Application-based or semantic constraints or business rules

DEFINITION

→ Some restrictions reflect complex **business rules**.

→ Due to their complexity can not be expressed in the data model, so that should be implemented at the level of code in the application or in a trigger.

Application-based or semantic constraints or business rules

DEFINITION

→ Some restrictions reflect complex **business rules**.

→ Due to their complexity can not be expressed in the data model, so that should be implemented at the level of code in the application or in a trigger.

EXAMPLE

The salary of an employee may not decrease, only increase

Schema-based constraints or explicit constraints

DEFINITION
→ Are defined when creating the schema in the deployment model (relational in this case).   → Are expressed in the language DDL

Schema-based constraints or explicit constraints

DEFINITION

➔ Are defined when creating the schema in the deployment model (relational in this case).

➔ Are expressed in the language DDL

CATEGORIES

➔ Domain Constraints.

➔ Key Constraints and Constraints on NULL Values.

➔ Integrity, Referential Integrity and Foreign Keys.

Domain Constraint

➔ Specify that the value of each attribute A must be an atomic value from the domain dom(A).

➔ The data types associated with the domains typically include standard numeric data types for integers.

➔ Other possible domains may be described by a subrange of values from a data type or as an enumerated data type in which all possible values are explicitly listed.

Key Constraints and Constraints on NULL Values

DEFINITION

➔ A relation is defined as a set of tuples.

➔ By definition all elements of a set are distinct; hence, all tuples in a relation must also be distinct.

Key Constraints and Constraints on NULL Values

DEFINITION

➜ A relation is defined as a set of tuples.

➜ By definition all elements of a set are distinct; hence, all tuples in a relation must also be distinct.

➜ Denote one subset of attributes by SK; then for any two distinct tuples t1 and t2 in a relation state r of R, we have the constraint that:

$$t_1[SK] \neq t_2[SK]$$

Key Constraints and Constraints on NULL Values

DEFINITION
➜ Any such set of attributes SK is called a superkey of the relation schema R.
➜ A superkey SK specifies a uniqueness constraint that no two distinct tuples in any state r of R can have the same value of SK.

Key Constraints and Constraints on NULL Values

➔ Any such set of attributes SK is called a superkey of the relation schema R.

➔ A superkey SK specifies a uniqueness constraint that no two distinct tuples in any state r of R can have the same value of SK.

➔ Every relation has at least one default superkey— the set of all its attributes.

➔ A key K of a relation schema R is a superkey of R with the additional property that removing any attribute A from K leaves a set of attributes K that is not a superkey of R any more.

Key Constraints and Constraints on NULL Values

DEFINITION

→ It is a minimal superkey—that is a superkey from which we cannot remove any attributes and still have the uniqueness constraint.

Key Constraints and Constraints on NULL Values

CURSO

COD_CURSO	GRUPO	PROFESOR	CUPO
CE-3101	1	1	30
CE-3101	2	2	40
CE-3002	1	3	10

SuperKey

CURSO

COD_CURSO	GRUPO	PROFESOR	CUPO
CE-3101	1	1	30
CE-3101	2	2	40
CE-3002	1	3	10

SuperKey

Key Constraints and Constraints on NULL Values

CURSO

COD_CURSO	GRUPO	PROFESOR	CUPO
CE-3101	1	1	30
CE-3101	2	2	40
CE-3002	1	3	10

Key

Key Constraints and Constraints on NULL Values

➔ A relation may have more than one key

CAR

License_number	Engine_serial_number	Make	Model	Year
Texas ABC-739	A69352	Ford	Mustang	02
Florida TVP-347	B43696	Oldsmobile	Cutlass	05
New York MPO-22	X83554	Oldsmobile	Delta	01
California 432-TFY	C43742	Mercedes	190-D	99
California RSK-629	Y82935	Toyota	Camry	04
Texas RSK-629	U028365	Jaguar	XJS	04

➔ The choice of one to become the primary key is somewhat arbitrary.

➔ It is usually better to choose a primary key with a single attribute or a small number of attributes.

Key Constraints and Constraints on NULL Values

➔ A relation may have more than one key

CAR

License_number	Engine_serial_number	Make	Mod
Texas ABC-739	A69352	Ford	Musta
Florida TVP-347	B43696	Oldsmobile	Cutlas
New York MPO-22	X83554	Oldsmobile	Delta
California 432-TFY	C43742	Mercedes	190-D
California RSK-629	Y82935	Toyota	Camr
Texas RSK-629	U028365	Jaguar	XJS

Attributes specifies whether NULL values are or are not permitted

➔ The choice of one to become the primary key arbitrary.

➔ It is usually better to choose a primary key with attribute or a small number of attributes.

Integrity

➜ The entity integrity constraint states that no primary key value can be NULL

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

Integrity

→ The entity integrity constraint states that no primary key value can be NULL

NULL values are not permitted

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

Referential Integrity

→ The referential integrity constraint is specified between two relations and is used to maintain the consistency among tuples in two relations.

→ Informally, the referential integrity constraint states that a tuple in one relation that refers to another relation must refer to an existing tuple in that relation.

Referential Integrity

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

Respues a pregunta inicial

Las jirafas no emiten sonido alguno convirtiéndose de esta manera en el único mamífero con esta característica

Relational Data Model

CE3101 - Bases de Datos