Mansoura University
Faculty of Computers and Information
Department of Information System
First Semester- 2020-2021

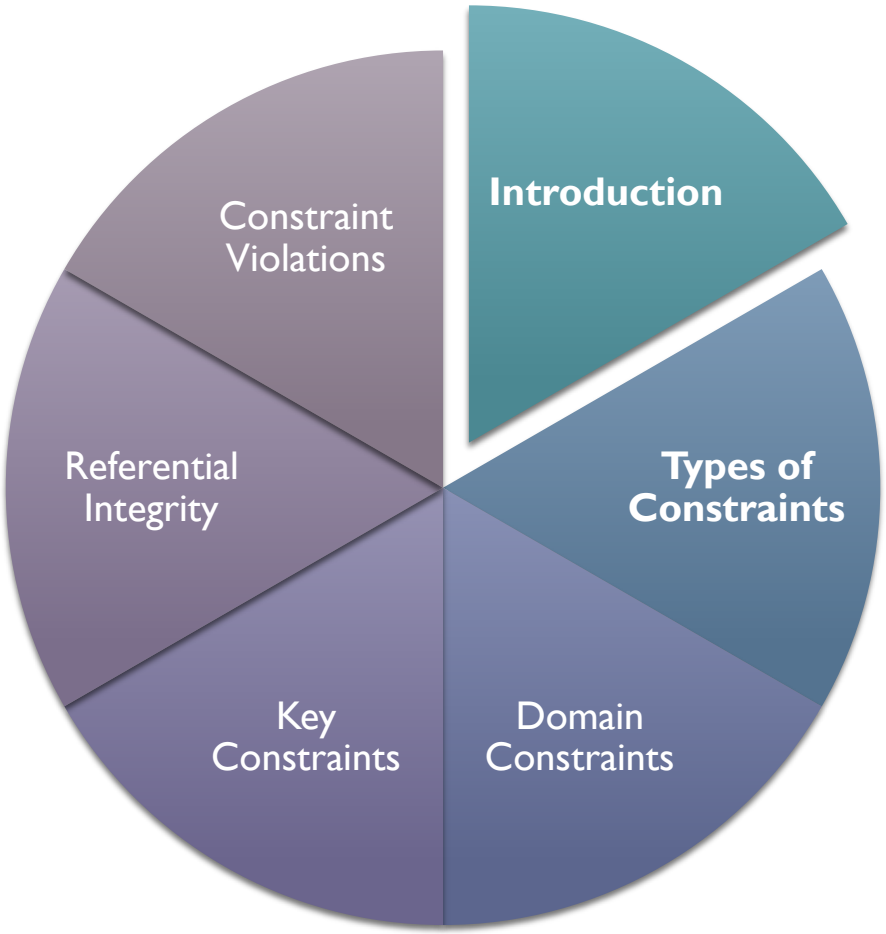
[IS313P] Database System II

Grade: 3 rd. IS & IT

Dr. Amira Rezk

INTEGRITY

AGENDA


Introduction
Types of Constraints
Domain Constraints
Key Constraints
Referential Integrity
Constraint Violations

INTRODUCTION

- Integrity is referring to the accuracy or correctness of the data in database.
- It should not be confused with security
- In both cases (security and integrity) the system needs to be aware of certain rules that users must not violate.
- This rules must be maintained in system catalog
- DBMS must monitor user operations to ensure that the rules are enforced.

INTRODUCTION

- An **integrity constraint** (constraint for short) is basically just a Boolean expression that must evaluate to TRUE.
- Constraints are so called because they constrain the values that can legally appear in some particular context.

INTRODUCTION

- Integrity constraints are specified and enforced at different times
  - 1. When the DBA or end user defines a database schema, he or she species the ICs that must hold on any instance of this database.
  - 2. When a database application is run, the DBMS checks for violations and disallows changes to the data that violate the specified ICs. (In some situations, rather than disallow the change, the DBMS might instead make some compensating changes to the data to ensure that the database instance satisfies all ICs.
  - In any case, changes to the database are not allowed to create an instance that violates any IC.)

TYPES OF CONSTRAINTS

- Constraints on databases can generally be divided into three main categories:
  - Constraints that are inherent in the data model. We call these inherent model-based constraints or implicit constraints.
  - Constraints that can be directly expressed in schemas of the data model, typically by specifying them in the DDL (data definition language, We call these schema-based constraints or explicit constraints.
  - Constraints that cannot be directly expressed in the schemas of the data model, and hence must be expressed and enforced by the application programs. We call these application-based or semantic constraints or business rules.

SCHEMA-BASED CONSTRAINTS
DOMAIN CONSTRAINTS

- Domain constraints specify that within each tuple, the value of each attribute A must be an atomic value from the domain dom(A).
- The data types associated with domains typically include:
  - Numeric data types for integers and real numbers
  - Characters
  - Booleans
  - Fixed-length strings
  - Variable-length strings
  - Date, time, timestamp
  - Money
  - Other special data types
- 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.

SCHEMA-BASED CONSTRAINTS
KEY AND NULL VALUES CONSTRAINTS

- No two tuples can have the same combination of values for all their attributes.
- Superkey
  - No two distinct tuples in any state r of R can have the same value for SK
- Key
  - Superkey of R
  - Removing any attribute A from K leaves a set of attributes K that is not a superkey of R any more
  - Key satisfies two properties:
    - Two distinct tuples in any state of relation cannot have identical values for (all) attributes in key
    - Minimal superkey
      - Cannot remove any attributes and still have uniqueness constraint in above condition hold

SCHEMA-BASED CONSTRAINTS
KEY AND NULL VALUES CONSTRAINTS (CONT.)

- Candidate key
  - Relation schema may have more than one key
- Primary key of the relation
  - Designated among candidate keys
  - Underline attribute
- Other candidate keys are designated as unique keys
- Another constraint on attributes specifies whether NULL values are or are not permitted.
- For example, if every STUDENT tuple must have a valid, non-NULL value for the Name attribute, then Name of STUDENT is constrained to be NOT NULL.

SCHEMA-BASED CONSTRAINTS
KEY AND NULL VALUES CONSTRAINTS (CONT.)

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

Figure 3.4
The CAR relation, with two candidate keys: License_number and Engine_serial_number.

RELATIONAL DATABASES AND RELATIONAL DATABASE SCHEMAS

- Relational database schema S
  - Set of relation schemas S = {R1, R2, ..., Rm}
  - Set of integrity constraints IC
- Relational database state
  - Set of relation states DB = {r1, r2, ..., rm}
  - Each ri is a state of Ri and such that the ri relation states satisfy integrity constraints specified in IC
- Invalid state
  - Does not obey all the integrity constraints
- Valid state
  - Satisfies all the constraints in the defined set of integrity constraints IC

INTEGRITY, REFERENTIAL INTEGRITY, AND FOREIGN KEYS

- Entity integrity constraint
  - No primary key value can be NULL
- Referential integrity constraint
  - Specified between two relations
  - Maintains consistency among tuples in two relations
- Foreign key rules:
  - The attributes in FK have the same domain(s) as the primary key attributes PK
  - Value of FK in a tuple t1 of the current state r1(R1) either occurs as a value of PK for some tuple t2 in the current state r2(R2) or is NULL

INTEGRITY, REFERENTIAL INTEGRITY, AND FOREIGN KEYS (CONT'D.)

- Diagrammatically display referential integrity constraints
  - Directed arc from each foreign key to the relation it references
- All integrity constraints should be specified on relational database schema

OTHER TYPES OF CONSTRAINTS

- Semantic integrity constraints
  - May have to be specified and enforced on a relational database
  - Use triggers and assertions
  - More common to check for these types of constraints within the application programs

OTHER TYPES OF CONSTRAINTS (CONT'D.)

- Functional dependency constraint
  - Establishes a functional relationship among two sets of attributes X and Y
  - Value of X determines a unique value of Y
- State constraints
  - Define the constraints that a valid state of the database must satisfy
- Transition constraints
  - Define to deal with state changes in the database

UPDATE OPERATIONS, TRANSACTIONS, AND DEALING WITH CONSTRAINT VIOLATIONS

- Operations of the relational model can be categorized into retrievals and updates
- Basic operations that change the states of relations in the database:
  - Insert
  - Delete
  - Update (or Modify)

Figure 3.6
One possible database state for the COMPANY relational database schema.

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

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS_ON

Essn	Pno	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

PROJECT

Pname	Pnumber	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

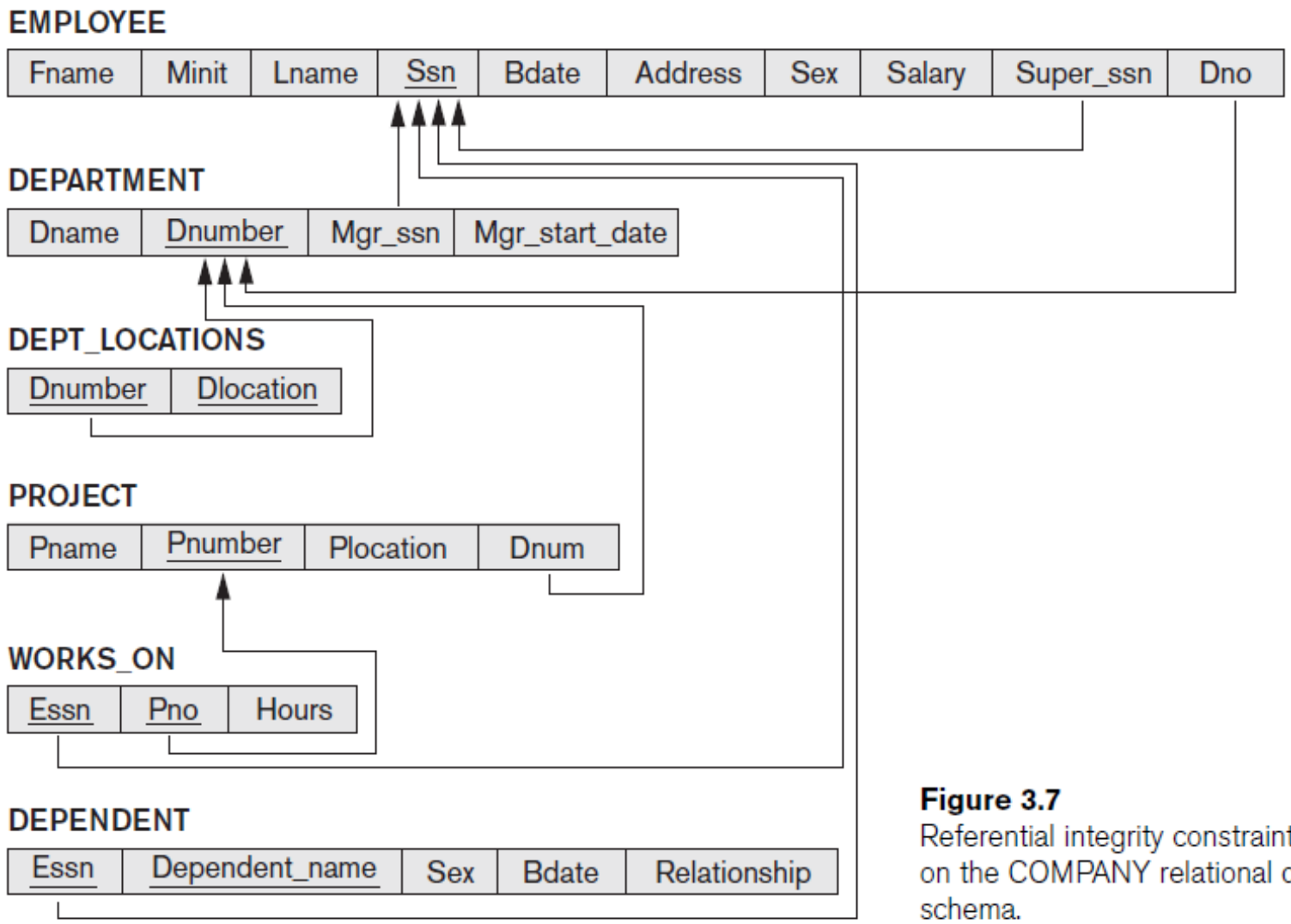


Figure 3.7
Referential integrity constraints displayed on the COMPANY relational database schema.

THE INSERT OPERATION

- Provides a list of attribute values for a new tuple t that is to be inserted into a relation R
- Can violate any of the four types of constraints
- If an insertion violates one or more constraints
  - Default option is to reject the insertion

THE DELETE OPERATION

- Can violate only referential integrity
  - If tuple being deleted is referenced by foreign keys from other tuples
  - Restrict
    - Reject the deletion
  - Cascade
    - Propagate the deletion by deleting tuples that reference the tuple that is being deleted
  - Set null or set default
    - Modify the referencing attribute values that cause the violation

THE UPDATE OPERATION

- Necessary to specify a condition on attributes of relation
  - Select the tuple (or tuples) to be modified
- If attribute not part of a primary key nor of a foreign key
  - Usually causes no problems
- Updating a primary/foreign key
  - Similar issues as with Insert/Delete

Questions?