Normalization:

- A method of splitting database tables into smaller tables to reduce redundancy and improve the integrity.

Types of Anomalies:

1. Update Anomaly
2. Insert Anomaly
3. Delete Anomaly

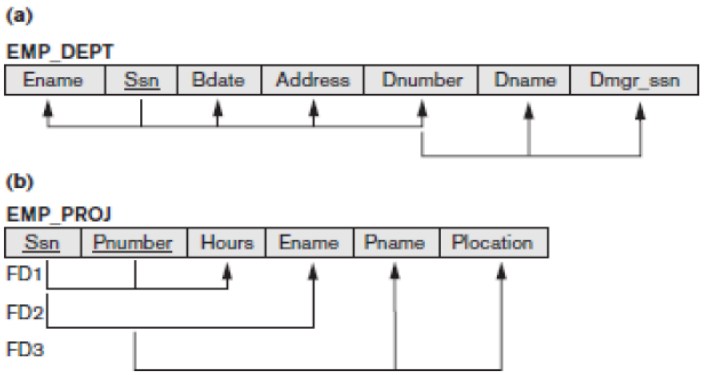
Normal Forms:

1. 1NF
   a. No multivalued attributes
   b. No composite attributes
   c. Always define a primary key
2. 2NF
   a. Must satisfy 1NF
   b. Each non-key attribute must depend on the entire primary key (no partial dependency)
3. 3NF
   a. Must satisfy 2NF
   b. A non-key attribute cannot depend on another non-key attribute (no transitive dependency)

Questions

15.20 What update anomalies occur in the EMP_PROJ and EMP_DEPT relations of Figure 15.3 and 15.4?

Figure 15.3
Two relation schemas suffering from update anomalies. (a) EMP_DEPT and (b) EMP_PROJ.



Redundancy

EMP_DEPT

Ename	Ssn	Bdate	Address	Dnumber	Dname	Dmgr_ssn
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Research	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Research	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administration	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administration	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Research	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Research	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administration	987654321
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Headquarters	888665555

Redundancy Redundancy

EMP_PROJ

Ssn	Pnumber	Hours	Ename	Pname	Plocation
123456789	1	32.5	Smith, John B.	ProductX	Bellaire
123456789	2	7.5	Smith, John B.	ProductY	Sugarland
666884444	3	40.0	Narayan, Ramesh K.	ProductZ	Houston
453453453	1	20.0	English, Joyce A.	ProductX	Bellaire
453453453	2	20.0	English, Joyce A.	ProductY	Sugarland
333445555	2	10.0	Wong, Franklin T.	ProductY	Sugarland
333445555	3	10.0	Wong, Franklin T.	ProductZ	Houston
333445555	10	10.0	Wong, Franklin T.	Computerization	Stafford
333445555	20	10.0	Wong, Franklin T.	Reorganization	Houston
999887777	30	30.0	Zelaya, Alicia J.	Newbenefits	Stafford
999887777	10	10.0	Zelaya, Alicia J.	Computerization	Stafford
987987987	10	35.0	Jabbar, Ahmad V.	Computerization	Stafford
987987987	30	5.0	Jabbar, Ahmad V.	Newbenefits	Stafford
987654321	30	20.0	Wallace, Jennifer S.	Newbenefits	Stafford
987654321	20	15.0	Wallace, Jennifer S.	Reorganization	Houston
888665555	20	Null	Borg, James E.	Reorganization	Houston

Figure 15.4
Sample states for EMP_DEPT and EMP_PROJ resulting from applying NATURAL JOIN to the relations in Figure 15.2. These may be stored as base relations for performance reasons.

15.29 Consider the following relations for an order-processing application database at ABC, Inc.

ORDER (O#, Odate, Cust#, Total_amount)
ORDER-ITEM (O#, I#, Qty_ordered, Total_price, Discount%)

Assume that each item has a different discount. The Total_price refers to one item, Odate is the date on which the order was placed, and the Total_amount is the amount of the order. If we apply a natural join on the relations Order-Item and Order in this database, what does the resulting relation schema look like? What will be its key? Show the FDs in this resulting relation. Is it in 2NF? Is it in 3NF? Why or why not? (State any assumptions you make.)

15.30 Consider the following relation:

CAR_SALE(**Car#,** Date_sold, **Salesman#**, Commission%, Discount_amt)

Assume that a car may be sold by multiple salesmen and hence {CAR#, SALESMAN#} is the primary key. Additional dependencies are:

Date_sold → Discount_amt and

Salesman# → commission% and

Car# → Date_sold

Based on the given primary key, is this relation in 1NF, 2NF, or 3NF? Why or why not? How would you successively normalize it completely?

15.31 Consider the following relation for published books:

BOOK (**Book_title, Authorname**, Book_type, Listprice, Author_affil, Publisher)

Author_affil referes to the affiliation of the author. Suppose the following dependencies exist:

Book_title → Publisher, Book_type

Book_type → List_price

Author_name → Author_affil

What normal form is the relation in? Explain your answer.

a. Apply normalization until you cannot decompose the relations further. State the reasons behind each decomposition.

15.32 This exercise asks you to convert business statements into dependencies. Consider the following relation

DISK_DRIVE (serialNumber, manufacturer, model, batch, capacity, retailer)

Each tuple in the relation DISK_DRIVE contains information about a disk drive with a unique serialNumber, made by a manufacturer, with a particular model, released in a certain batch, which has a certain storage capacity, and is sold by a certain retailer. For example, the tuple DISK_DRIVE (1978619, WesternDigital, A2235X, 765234, 500, CompUSA) specifies that WesternDigital made a disk drive with serial number 1978619, model number A2235X in batch 765235 with 500GB that is sold by CompUSA.

Write each of the following dependencies as an FD:

a. The manufacturer and serial number uniquely identifies the drive .

b. A model number is registered by a manufacturer and hence can't be used by another manufacturer.
c. All disk drives in a particular batch are the same model.
d. All disk drives of a particular model of a particular manufacturer have exactly the same capacity.

15.33 Consider the following relation:

R (Doctor#, Patient#, Date, Diagnosis, Treat_code, Charge)

In this relation, a tuple describes a visit of a patient to a doctor along with a treatment code and daily charge. Assume that diagnosis is determined (uniquely) for each patient by a doctor. Assume that each treatment code has a fixed charge (regardless of patient). Is this relation in 2NF? Justify your answer and decompose if necessary. Then argue whether further normalization to 3NF is necessary, and if so, perform it.

No

Doctor # Patient # Diagnosis