

F.K.

<u>Ename</u>	<u>Ssn</u>	Bdate	Address	<u>Dnumber</u>
--------------	------------	-------	---------	----------------

P.K.

Figure 1: **EMPLOYEE**

F.K.

Dname	<u>Dnumber</u>	<u>Dmgr_ssn</u>
-------	----------------	-----------------

P.K.

Figure 2: **DEPARTMENT**

F.K.

<u>Dnumber</u>	<u>Dlocation</u>
----------------	------------------

P.K.

Figure 3: **DEPT\_LOCATIONS**

F.K.

Pname	<u>Pnumber</u>	Plocation	<u>Dnum</u>
-------	----------------	-----------	-------------

P.K.

Figure 4: **PROJECT**

F.K.      F.K.

<u>Ssn</u>	<u>Pnumber</u>	Hours
------------	----------------	-------

P.K.

Figure 5: **WORKS\_ON**

**EMPLOYEE**

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

**DEPARTMENT**

Dname	<u>Dnumber</u>	<u>Dmgr_ssn</u>
Research	5	333445555
Administration	4	987654321
Headquarters	1	888665555

**DEPT\_LOCATIONS**

<u>Dnumber</u>	<u>Dlocation</u>
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

**WORKS\_ON**

<u>Ssn</u>	<u>Pnumber</u>	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	<u>Pnumber</u>	Plocation	<u>Dnum</u>
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

Figure 6: Sample database for state for the relational database schema in Figures 1 – 5.

Ename	<u>Ssn</u>	Bdate	Address	Dnumber	Dname	Dmgr_ssn
-------	------------	-------	---------	---------	-------	----------

Figure 7: **EMP\_DEPT** (EMPLOYEE  $\bowtie$  DEPARTMENT)

<u>Ssn</u>	<u>Pnumber</u>	Hours	Ename	Pname	Plocation
------------	----------------	-------	-------	-------	-----------

Figure 8: **EMP\_PROJ** (II...(EMPLOYEE  $\bowtie$  PROJECT  $\bowtie$  WORKS\_ON))

EMP_DEPT							Redundancy
Ename	<u>Ssn</u>	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	

EMP_PROJ						Redundancy	Redundancy
<u>Ssn</u>	<u>Pnumber</u>	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 9: Sample database for state for the relational database schema in Figures 7 – 8.

<u>Ename</u>	<u>Plocation</u>
--------------	------------------

Figure 10: **EMP\_LOCS** ( $\Pi_{\dots}(\text{EMP\_PROJ})$ )

<u>Ssn</u>	<u>Pnumber</u>	Hours	Pname	Plocation
------------	----------------	-------	-------	-----------

Figure 11: **EMP\_PROJ1** ( $\Pi_{\dots}(\text{EMP\_PROJ})$ )

**EMP\_LOCS**

Ename	Plocation
Smith, John B.	Bellaire
Smith, John B.	Sugarland
Narayan, Ramesh K.	Houston
English, Joyce A.	Bellaire
English, Joyce A.	Sugarland
Wong, Franklin T.	Sugarland
Wong, Franklin T.	Houston
Wong, Franklin T.	Stafford
Zelaya, Alicia J.	Stafford
Jabbar, Ahmad V.	Stafford
Wallace, Jennifer S.	Stafford
Wallace, Jennifer S.	Houston
Borg, James E.	Houston

**EMP\_PROJ1**

Ssn	Pnumber	Hours	Pname	Plocation
123456789	1	32.5	ProductX	Bellaire
123456789	2	7.5	ProductY	Sugarland
666884444	3	40.0	ProductZ	Houston
453453453	1	20.0	ProductX	Bellaire
453453453	2	20.0	ProductY	Sugarland
333445555	2	10.0	ProductY	Sugarland
333445555	3	10.0	ProductZ	Houston
333445555	10	10.0	Computerization	Stafford
333445555	20	10.0	Reorganization	Houston
999887777	30	30.0	Newbenefits	Stafford
999887777	10	10.0	Computerization	Stafford
987987987	10	35.0	Computerization	Stafford
987987987	30	5.0	Newbenefits	Stafford
987654321	30	20.0	Newbenefits	Stafford
987654321	20	15.0	Reorganization	Houston
888665555	20	NULL	Reorganization	Houston

Figure 12: Sample database for state for the relational database schema in Figures 10 – 11.

Ssn	Pnumber	Hours	Pname	Plocation	Ename
123456789	1	32.5	ProductX	Bellaire	Smith, John B.
*	123456789	1	32.5	ProductX	English, Joyce A.
	123456789	2	7.5	ProductY	Smith, John B.
*	123456789	2	7.5	ProductY	English, Joyce A.
*	123456789	2	7.5	ProductY	Wong, Franklin T.
	666884444	3	40.0	ProductZ	Narayan, Ramesh K.
*	666884444	3	40.0	ProductZ	Wong, Franklin T.
*	453453453	1	20.0	ProductX	Smith, John B.
	453453453	1	20.0	ProductX	English, Joyce A.
*	453453453	2	20.0	ProductY	Smith, John B.
	453453453	2	20.0	ProductY	English, Joyce A.
*	453453453	2	20.0	ProductY	Wong, Franklin T.
*	333445555	2	10.0	ProductY	Smith, John B.
*	333445555	2	10.0	ProductY	English, Joyce A.
	333445555	2	10.0	ProductY	Wong, Franklin T.
*	333445555	3	10.0	ProductZ	Narayan, Ramesh K.
	333445555	3	10.0	ProductZ	Wong, Franklin T.
	333445555	10	10.0	Computerization	Wong, Franklin T.
*	333445555	20	10.0	Reorganization	Narayan, Ramesh K.
	333445555	20	10.0	Reorganization	Wong, Franklin T.

\*  
\*  
\*

Figure 13: Result of applying NATURAL JOIN to the tuples in **EMP\_PROJ1** and **EMP\_LOCS** of Figure 12.

## Informal Design Guidelines for Relation Schemas

1. Making sure that the semantics of the attributes is clear in the schema
  2. Reducing the redundant information in tuples
  3. Reducing the NULL values in tuples
  4. Disallowing the possibility of generating spurious tuples
1. **Imparting Clear Semantics to Attributes in Relations:** Whenever we group attributes to form a relation schema, we assume that attributes belonging to one relation have certain real-world meaning and a proper interpretation associated with them. The **semantics** of a relation refers to its meaning resulting from the interpretation of attribute values in a tuple. The ease with which the meaning of a relation's attributes can be explained is an informal measure of how well the relation is designed.

Design a relation schema so that it is easy to explain its meaning. Do not combine attributes from multiple entity types and relationship types into a single relation. Intuitively, if a relation schema corresponds to one entity type or one relationship type, it is straightforward to explain its meaning. Otherwise, if the relation corresponds to a mixture of multiple entities and relationships, semantic ambiguities will result and the relation cannot be easily explained.

2. **Redundant Information in Tuples and Update Anomalies:** One goal of schema design is to minimize the storage space used by the base relations (and hence the corresponding files). Grouping attributes into relation schemas has a significant effect on storage space.

Storing **natural joins** of base relations leads to an additional problem referred to as **update anomalies**. These can be classified into insertion anomalies, deletion anomalies, and modification anomalies. It is easy to see that these three anomalies are undesirable and cause difficulties to maintain consistency of data as well as require unnecessary updates that can be avoided.

Design the base relation schemas so that no insertion, deletion, or modification anomalies are present in the relations. If any anomalies are present, note them clearly and make sure that the programs that update the database will operate correctly.

3. **NULL Values in Tuples:** In some schema designs we may group many attributes together into a “fat” relation. If many of the attributes do not apply to all tuples in the relation, we end up with many NULLs in those tuples.

- Space wastage at the storage level
- Problems with understanding the meaning of the attributes
- Handling of aggregate operations such as COUNT or SUM
- Problem with comparisons as we may get UNKNOWN

As far as possible, avoid placing attributes in a base relation whose values may frequently be NULL. If NULLs are unavoidable, make sure that they apply in exceptional cases only and do not apply to a majority of tuples in the relation.

Using space efficiently and avoiding joins with NULL values are the two overriding criteria that determine whether to include the columns that may have NULLs in a relation or to have a separate relation for those columns (with the appropriate key columns).

4. **Generation of Spurious Tuples:**

Design relation schemas so that they can be joined with equality conditions on attributes that are appropriately related (primary key, foreign key) pairs in a way that guarantees that no spurious tuples are generated. Avoid relations that contain matching attributes that are not (foreign key, primary key) combinations because joining on such attributes may produce spurious tuples.