Relations and Set Properties
a RELATION is a SET

- With All Set Properties
- No Ordering Among Tuples
- No Duplicate Tuples in Set
- Each Tuple is Unique
Uniqueness in Databases

• Basic concept in DB:
  – I give DB a little info, like SSN
  – DB gives me back a lot of information:
    • Name, address, salary, department etc.
• Not only is every tuple unique
• Among all the tuples some field(s) is/are unique
Keys and Superkeys

• Aim: field or smallest set of fields that will give uniqueness
• Starting point: the set of all fields is guaranteed unique
  – Because no duplicates
• Defining key: start with set of all fields: superkey.
Superkeys

• A Subset of the ATTRIBUTES which distinguishes among tuples is the SUPERKEY

• Example: (FNAME, LNAME, SSN) Distinguishes Employee Tuples
  – No two employees have the same (FNAME, LNAME, SSN)
Keys and Superkeys

- If it has **no unneeded attributes**, SUPERKEY is a KEY
- Example: SSN alone Distinguishes Employees. FNAME, LNAME not needed
Every relation has a key

- Sometimes all attributes
- Often a subset

<table>
<thead>
<tr>
<th>Dnumber</th>
<th>Dept_loc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bellaire</td>
</tr>
<tr>
<td>4</td>
<td>Sugarland</td>
</tr>
<tr>
<td>5</td>
<td>Houston</td>
</tr>
<tr>
<td>5</td>
<td>Bellair</td>
</tr>
<tr>
<td>5</td>
<td>Sugarland</td>
</tr>
</tbody>
</table>

DEPT_LOCATIONS is ALL KEY
More Than One Key

• Sometimes 2 keys exist
• Each is a CANDIDATE KEY
• one chosen by DB Designer as PRIMARY KEY
• PRIMARY KEY always exists
CANDIDATE KEYS

<table>
<thead>
<tr>
<th>EMPLOYEE</th>
<th>FNAME</th>
<th>LNAME</th>
<th>SSN</th>
<th>SAL</th>
<th>SUPERSSN</th>
<th>DNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>Smith</td>
<td></td>
<td>123456789</td>
<td>30K</td>
<td>333445555</td>
<td>5</td>
</tr>
<tr>
<td>Frank</td>
<td>Wong</td>
<td></td>
<td>333445555</td>
<td>40K</td>
<td>888665555</td>
<td>5</td>
</tr>
<tr>
<td>Alice</td>
<td>Zelaya</td>
<td></td>
<td>999887777</td>
<td>25K</td>
<td>987654321</td>
<td>4</td>
</tr>
<tr>
<td>Jenny</td>
<td>Wallace</td>
<td></td>
<td>987654321</td>
<td>43K</td>
<td>888665555</td>
<td>4</td>
</tr>
<tr>
<td>Ram</td>
<td>Narayan</td>
<td></td>
<td>666884444</td>
<td>38K</td>
<td>333445555</td>
<td>5</td>
</tr>
<tr>
<td>Joyce</td>
<td>English</td>
<td></td>
<td>453453453</td>
<td>25K</td>
<td>333445555</td>
<td>5</td>
</tr>
<tr>
<td>Ahmad</td>
<td>Jabbar</td>
<td></td>
<td>987987987</td>
<td>25K</td>
<td>987654321</td>
<td>4</td>
</tr>
<tr>
<td>James</td>
<td>Borg</td>
<td></td>
<td>888665555</td>
<td>55K</td>
<td>null</td>
<td>1</td>
</tr>
</tbody>
</table>

ONE CANDIDATE KEY

ANOTHER

ONE CHOSEN

Now this DB has a RULE: NO TWO EMPLOYEES with SAME SOC SEC NUMBER allowed in company.
Note of Reality

• In most database management systems it is in fact possible to have relations without a primary key and relations with duplicate tuples.

• Here we are talking about well-designed databases.