COS221 - L22 - Normalisation - 1NF, 2NF and 3NF

Linda Marshall

24 April 2023

First Normal Form - 1NF

- **First normal form** - All attributes of the relation must be atomic

First Normal Form - 1NF

Consider the following relation (a) and state (b):
Ignore Fig 14-9(c) for now

(a)

DEPARTMENT

Dname	Dnumber	Dmgr_ssn	Dlocations

(b)

DEPARTMENT

Dname	Dnumber	Dmgr_ssn	Dlocations
Research	5	333445555	{Bellaire, Sugarland, Houston}
Administration	4	987654321	{Stafford}
Headquarters	1	888665555	{Houston}

(c)

DEPARTMENT

Dname	Dnumber	Dmgr_ssn	Dlocation
Research	5	333445555	Bellaire
Research	5	333445555	Sugarland
Research	5	333445555	Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston

Figure 14.9
Normalization into 1NF. (a) A relation schema that is not in 1NF. (b) Sample state of relation DEPARTMENT. (c) 1NF version of the same relation with redundancy.

The relation is not in a normal form due to the **multivalued attribute** `Dlocations`

First Normal Form - 1NF

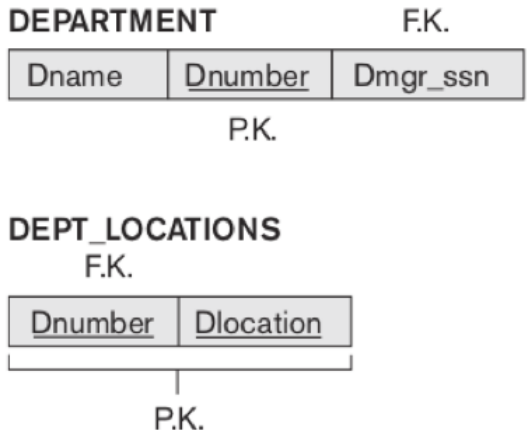
Follow one of the following to remove the multivalued attribute.

1. Remove the attribute that violates the atomic requirement. Create a new relation with the primary key of the original relation and the non-atomic attribute as the primary key.
2. Expand the key of the relation to include the non-atomic attribute as part of the key.
3. If the number of instances of the non-atomic attribute is known, replace the attribute with that many additional attributes.

First Normal Form - 1NF (Option 1)

Follow one of the following to remove the multivalued attribute.

1. **Remove the attribute that violates the atomic requirement. Create a new relation with the primary key of the original relation and the non-atomic attribute as the primary key.**



2. Expand the key of the relation to include the non-atomic attribute as part of the key.
3. If the number of instances of the non-atomic attribute is known, replace the attribute with that many additional attributes.

First Normal Form - 1NF (Option 2)

Follow one of the following to remove the multivalued attribute.

1. Remove the attribute that violates the atomic requirement. Create a new relation with the primary key of the original relation and the non-atomic attribute as the primary key.
2. **Expand the key of the relation to include the non-atomic attribute as part of the key.**

(c)

DEPARTMENT

Dname	Dnumber	Dmgr_ssn	Dlocation
Research	5	333445555	Bellaire
Research	5	333445555	Sugarland
Research	5	333445555	Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston

3. If the number of instances of the non-atomic attribute is known, replace the attribute with that many additional attributes.

First Normal Form - 1NF (Option 3)

Follow one of the following to remove the multivalued attribute.

1. Remove the attribute that violates the atomic requirement. Create a new relation with the primary key of the original relation and the non-atomic attribute as the primary key.
2. Expand the key of the relation to include the non-atomic attribute as part of the key.
3. **If the number of instances of the non-atomic attribute is known, replace the attribute with that many additional attributes.**
   `DEPARTMENT(Dname,Dnumber,Dmgr_ssn,Dlocation1, Dlocation2,Dlocation3)`

First Normal Form - 1NF

What about nested relations?

- Remove the nested relation attributes into a new relation and propagate the primary key to it. The primary key of the new relation will combine the partial key with the primary key of the original relation.

Consider the relation:
EMP_PROJS(Ssn, Ename, PROJS(Pnumber,Hours)) In this case, PROJS is nested and multivalued. Ssn is the primary key and Pnumber a partial key.

First Normal Form - 1NF

(a)

EMP_PROJ

		Projs	
Ssn	Ename	Pnumber	Hours

(b)

EMP_PROJ

Ssn	Ename	Pnumber	Hours
123456789	Smith, John B.	1	32.5
		2	7.5
666884444	Narayan, Ramesh K.	3	40.0
453453453	English, Joyce A.	1	20.0
		2	20.0
333445555	Wong, Franklin T.	2	10.0
		3	10.0
		10	10.0
		20	10.0
999887777	Zelaya, Alicia J.	30	30.0
		10	10.0
987987987	Jabbar, Ahmad V.	10	35.0
		30	5.0
987654321	Wallace, Jennifer S.	30	20.0
		20	15.0
888665555	Borg, James E.	20	NULL

(c)

EMP_PROJ1

Ssn	Ename

EMP_PROJ2

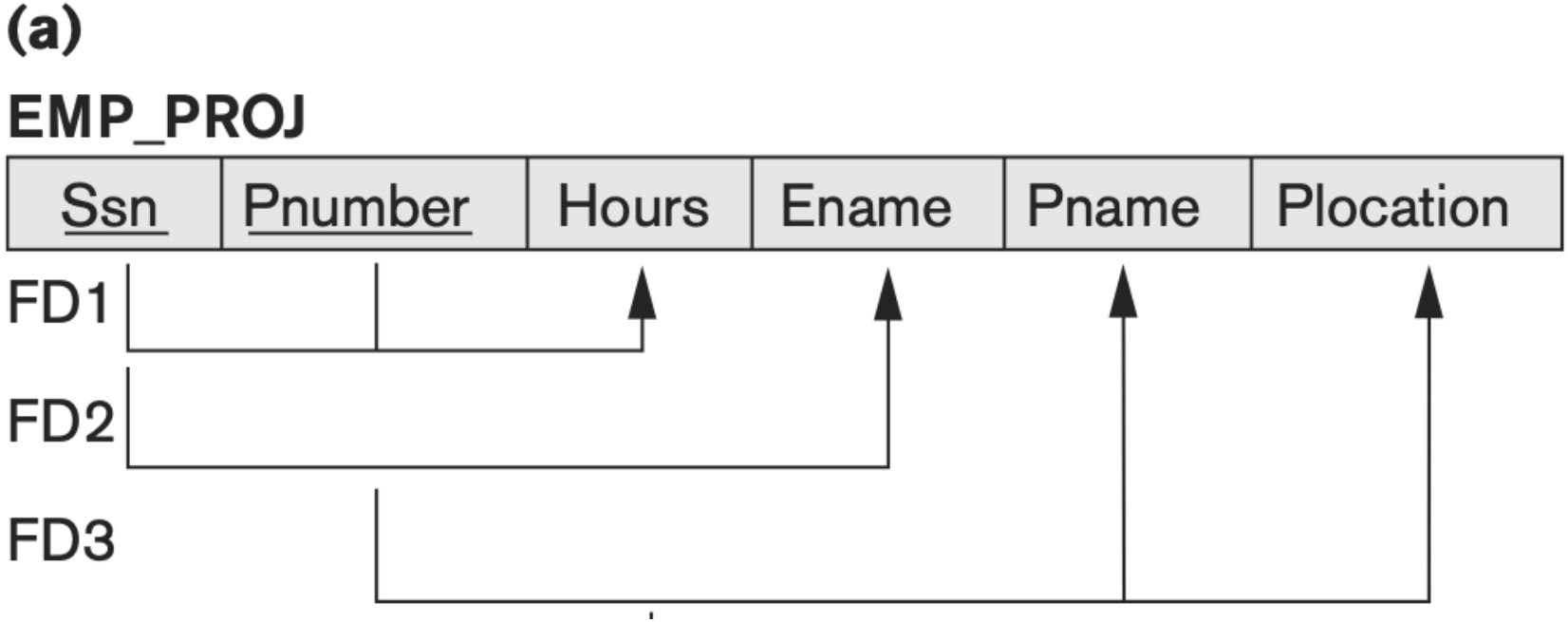
Ssn	Pnumber	Hours

Figure 14.10
Normalizing nested relations into 1NF. (a) Schema of the EMP_PROJ relation with a *nested relation* attribute PROJS. (b) Sample extension of the EMP_PROJ relation showing nested relations within each tuple. (c) Decomposition of EMP_PROJ into relations EMP_PROJ1 and EMP_PROJ2 by propagating the primary key.

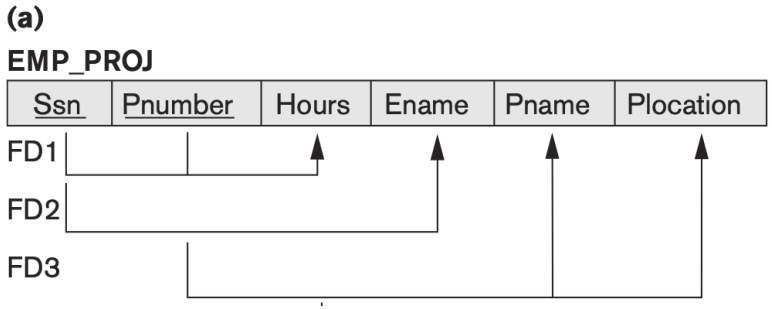
Second Normal Form - 2NF

- **Second normal form** - A relational schema R is in 2NF if it is in 1NF and if every nonprime attribute A in R is fully functionally dependent on the primary key of R.

Consider the following relation:



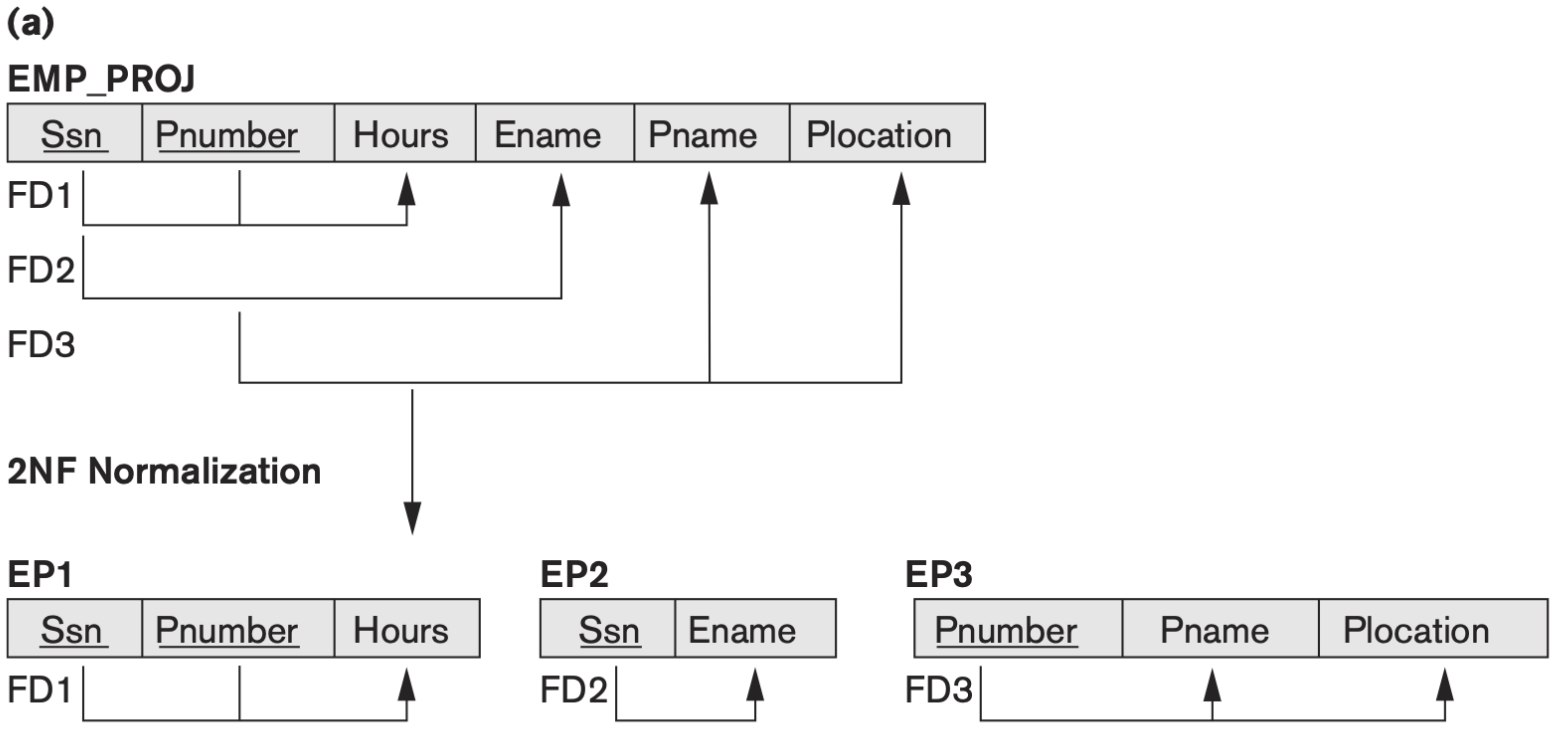
Second Normal Form - 2NF



- `Hours` is fully functionally dependent on the primary key (FD1)
- `Ename` is partially dependent on the primary key (FD2)
- `Pname` and `Plocation` are partially dependent on the primary key (FD3)

Second Normal Form - 2NF

- Remove `Ename` from EMP_PROJS into a new relation. Add the left side of the FD as its primary key
- Remove `Pname` and `Plocation` from EMP_PROJS into a new relation. Add the left side of the FD as its primary key.



Third Normal Form - 3NF

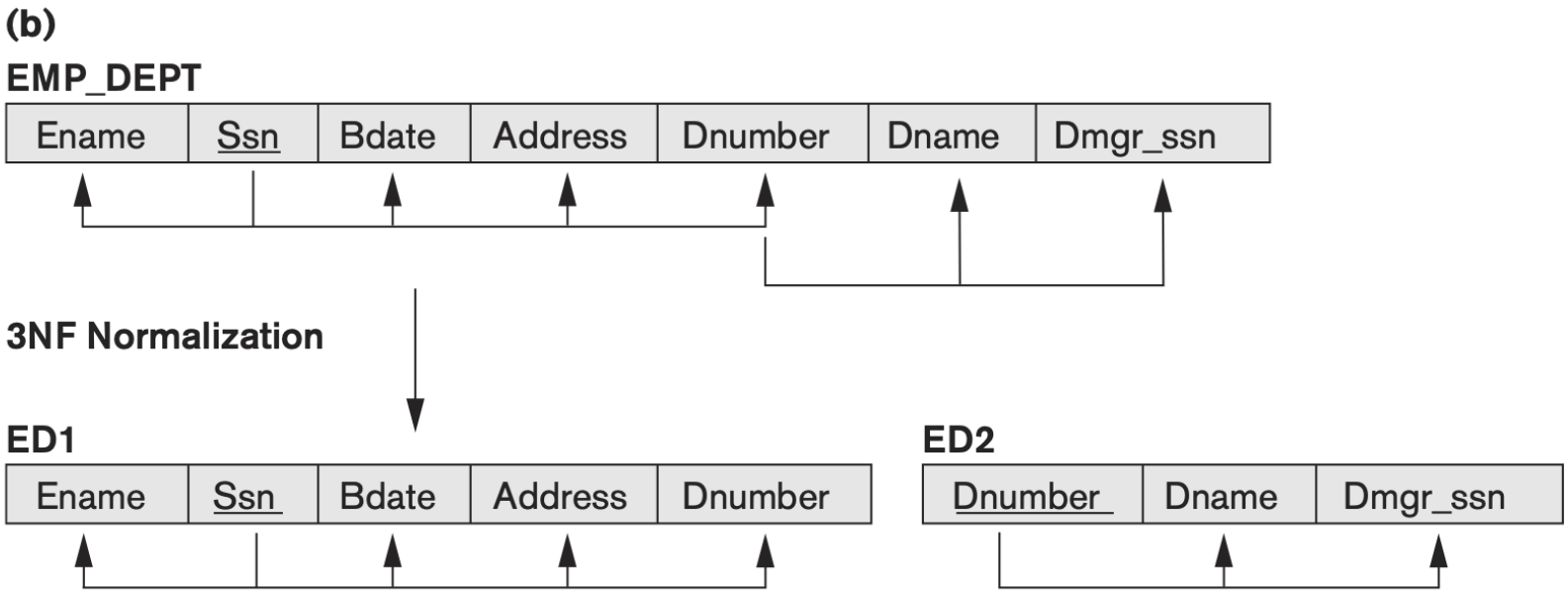
- **Third normal form** - A relation schema R is in 3NF if it is in 2NF and no nonprime attribute is transitively dependent on the primary key

Note: A functional dependency, $X \rightarrow Y$, in a relation schema R is a *transitive dependency* if there exists a set of attributes Z in R, that is neither a candidate key or a subset of any key or R and both $X \rightarrow Z$ and $Z \rightarrow Y$ hold.

Third Normal Form - 3NF

- Remove the attributes which are dependent on the attributes resulting in the transitive dependency (Z) and place them in a new relation. Make the attributes corresponding to Z in the transitive dependency the primary key of this new relation.
- Only attributes of Z should remain in the original relation.

Consider the following example. EMP_DEPT is in both 1NF and 2NF, but not in 3NF.



Summary of 1NF, 2NF and 3NF

Table 14.1 Summary of Normal Forms Based on Primary Keys and Corresponding Normalization

Normal Form	Test	Remedy (Normalization)
First (1NF)	Relation should have no multivalued attributes or nested relations.	Form new relations for each multivalued attribute or nested relation.
Second (2NF)	For relations where primary key contains multiple attributes, no nonkey attribute should be functionally dependent on a part of the primary key.	Decompose and set up a new relation for each partial key with its dependent attribute(s). Make sure to keep a relation with the original primary key and any attributes that are fully functionally dependent on it.
Third (3NF)	Relation should not have a nonkey attribute functionally determined by another nonkey attribute (or by a set of nonkey attributes). That is, there should be no transitive dependency of a nonkey attribute on the primary key.	Decompose and set up a relation that includes the nonkey attribute(s) that functionally determine(s) other nonkey attribute(s).

Normalise to 3NF

1. R(<u>a</u>, b, c, R1(<u>d</u>, <u>e</u>)) with FD's:
   $a \rightarrow \{b, c\}$
   $c \rightarrow \{d, e\}$
2. R(<u>a</u>, <u>b</u>, c, d, e, {f}) with FD's:
   $\{a,b\} \rightarrow c$
   $a \rightarrow d$
   $c \rightarrow e$
3. R(<u>a</u>, <u>b</u>, c, d, e, {f}) with FD's:
   $b \rightarrow c$
   $a \rightarrow d$
   $c \rightarrow e$