

Presentation on Functional Dependencies CS x265

This presentation assumes that you have previously viewed

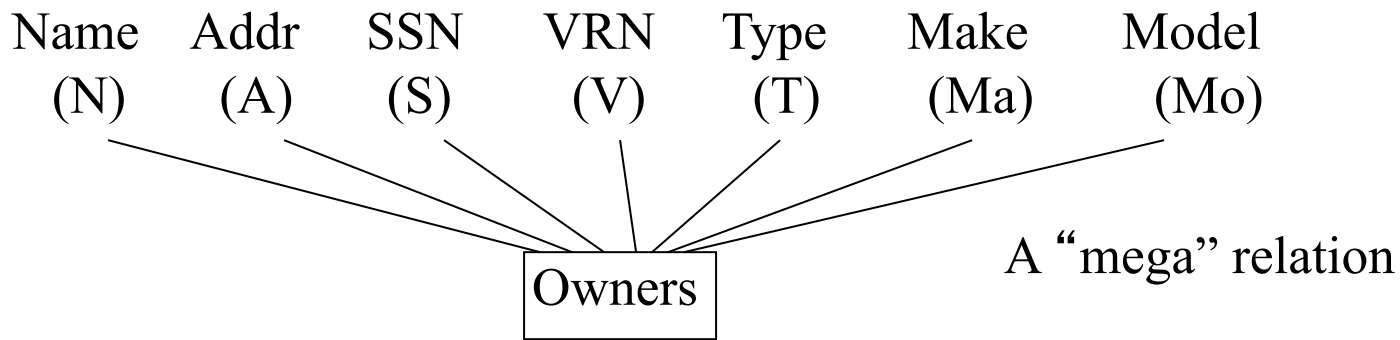
Watch videos (90 min) and answer questions from DB8

Relational Design Theory

(https://class.stanford.edu/courses/DB/RD/SelfPaced/courseware/ch-relational_design_theory/)

and/or read Chapter 3 of Ullman and Widom, Introduction to Database Management Systems

Assume the following relational schema covering vehicle ownership data (forgive lack of key, or assume that all attributes form the key, for now).



Individual persons, uniquely identified by SSN, are stored with their Name and Addr(ess), and are stored with information of the vehicles they own, where each vehicle is uniquely Identified by a Vehicle Registration Number (VRN), its Type (auto, truck, motorcycle), Manufacturer (aka Make), and Model. A sample of a database fitting this schema is below.

<u>N</u>	<u>A</u>	<u>S</u>	<u>V</u>	<u>T</u>	<u>Ma</u>	<u>Mo</u>
Fred	Nashville	123	987	Truck	Ford	Ranger
Sri	NewYork	234	876	Car	Toyota	Camry
Gabriel	Nashville	345	765	MotorCy	Harley	Hog
Fred	Nashville	123	654	Car	VW	Bug

Given your domain knowledge of vehicle ownership relationships in the real world, list functional dependencies that you believe should be asserted as true/required of this relational schema (and be enforced in the database).

Assume the following functional dependencies apply to the schema

$S \rightarrow N$, $S \rightarrow A$, $V \rightarrow T$, $V \rightarrow Ma$, $V \rightarrow Mo$, $Mo \rightarrow Ma$, $Mo \rightarrow T$, $V \rightarrow S$

Name	Addr	SSN	VRN	Type	Make	Model
(N)	(A)	(S)	(V)	(T)	(Ma)	(Mo)

State the constraint implied by each FD in English.

N	A	S	V	T	Ma	Mo	
Fred	Nashville	123	987	Truck	Ford	Ranger	
Sri	NewYork	234	876	Car	Toyota	Camry	
Gabriel	Nashville	345	765	MotorCy	Harley	Hog	
Fred	Nashville	123	654	Car	VW	Bug	
Sri	NewYork	234	654	Car	VW	Bug	*
Mary	LosAngeles	456	876	Car	Toyota	Corolla	*
Fred	Nshville	123	543	Car	Honda	Accord	*

Each of the final three rows cause a violation of at least one FD (given the earlier rows). Identify the FDs that are violated in each case.

Assume the following functional dependencies apply to the schema

$S \rightarrow N$, $S \rightarrow A$, $V \rightarrow T$, $V \rightarrow Ma$, $V \rightarrow Mo$, $Mo \rightarrow Ma$, $Mo \rightarrow T$, $V \rightarrow S$

Each vehicle associated with one owner (1..1) or at most one (0..1). Nothing about these FDs implies that a Person need be an owner (0..*)

Name (N)	Addr (A)	SSN (S)	VRN (V)	Type (T)	Make (Ma)	Model (Mo)	
Fred	Nashville	123	987	Truck	Ford	Ranger	
Sri	NewYork	234	876	Car	Toyota	Camry	
Gabriel	Nashville	345	765	MotorCy	Harley	Hog	
Fred	Nashville	123	654	Car	VW	Bug	
Sri	NewYork	234	654	Car	VW	Bug	($V \rightarrow S$)
Mary	LosAngeles	456	876	Car	Toyota	Corolla	($V \rightarrow Mo, V \rightarrow S$)
Fred	Nshville	123	543	Car	Honda	Accord	($S \rightarrow A$)

Of the FDs given, which are redundant (and not needed because they can be inferred from the remaining FDs)?

Give the key(s) of the relation above, as dictated by the FDs.

Are the FDs minimal? Can we infer an FD from the other FDs?

$S \rightarrow N, S \rightarrow A, V \rightarrow T, V \rightarrow Ma, V \rightarrow Mo, Mo \rightarrow Ma, Mo \rightarrow T, V \rightarrow S$

From $V \rightarrow Mo$ and $Mo \rightarrow Ma$ we know $V \rightarrow Ma$ ($V \rightarrow Ma$ not needed)

From $V \rightarrow Mo$ and $Mo \rightarrow T$ we know $V \rightarrow T$ ($V \rightarrow T$ not needed)

A minimal set of FDs (not necessarily unique set):

$S \rightarrow N, S \rightarrow A, \cancel{V \rightarrow T}, \cancel{V \rightarrow Ma}, V \rightarrow Mo, Mo \rightarrow Ma, Mo \rightarrow T, V \rightarrow S$

Systematic algorithm for determining minimal set of FDs (and ...) uses the *Attribute Closure algorithm* (if you know attribute A, what other attributes, B, can be determined)

Attribute closure of N is $\{N\}$ N not on LHS of any FD

Attribute closure of A is $\{A\}$

Attribute closure of S is $\{S\}$

Attribute closure of T is $\{T\}$

Attribute closure of Ma is $\{Ma\}$

Attribute closure of Mo is $\{Mo\}$

$S \xrightarrow{S \rightarrow N} \{S, N\} \xrightarrow{S \rightarrow A} \{S, N, A\}$

$Mo \xrightarrow{Mo \rightarrow T} \{Mo, T\} \xrightarrow{Mo \rightarrow Ma} \{Mo, T, Ma\}$

Extend the set using any FD with a LHS that is member of current set

$S \rightarrow N, S \rightarrow A, V \rightarrow T, V \rightarrow Ma, V \rightarrow Mo, Mo \rightarrow Ma, Mo \rightarrow T, V \rightarrow S$

Attribute closure of V is $\{V\} \xrightarrow{V \rightarrow Mo} \{V, Mo\} \xrightarrow{Mo \rightarrow Ma} \{V, Mo, Ma\}$

$Mo \rightarrow T \rightarrow \{V, Mo, Ma, T\} \xrightarrow{V \rightarrow S} \{V, Mo, Ma, T, S\}$

$S \rightarrow N \rightarrow \{V, Mo, Ma, T, S, N\} \xrightarrow{S \rightarrow A} \underline{\{V, Mo, Ma, T, S, N, A\}}$

Note that the attribute closure of V includes ALL attributes. **V is a key.**

In general, a key is any minimal set of attributes with attribute closures whose union includes all attributes. Only the attribute closure of V contains V, so V is the only key (though there are many super keys).

If V were not an attribute (and all FDs involving V were removed), then S, Mo would be the only key. In general, however, there may be more than one Key.

$S \rightarrow N, S \rightarrow A, V \rightarrow T, V \rightarrow Ma, V \rightarrow Mo, Mo \rightarrow Ma, Mo \rightarrow T, V \rightarrow S$

Attribute closure of V is $\{V\} \xrightarrow{V \rightarrow Mo} \{V, Mo\} \xrightarrow{Mo \rightarrow Ma} \{V, Mo, Ma\}$

$Mo \rightarrow T \rightarrow \{V, Mo, Ma, T\}$ $V \rightarrow S \rightarrow \{V, Mo, Ma, T, S\}$

$S \rightarrow N \rightarrow \{V, Mo, Ma, T, S, N\}$ $S \rightarrow A \rightarrow \underline{\{V, Mo, Ma, T, S, N, A\}}$

Incomplete Alg. for determining minimal FD set

Give an ordering of FDs (different orderings may lead to different minimal FD sets)

For each FD, $A \rightarrow B$,

does attribute closure of A include B when $A \rightarrow B$ is excluded from derivation?

if so, eliminate $A \rightarrow B$ and continue

$S \rightarrow N, S \rightarrow A, V \rightarrow T, V \rightarrow Ma, V \rightarrow Mo, Mo \rightarrow Ma, Mo \rightarrow T, V \rightarrow S$

Incomplete Alg. for determining minimal FD set

Give an ordering of FDs (different orderings may lead to different minimal FD sets)

For each FD, $A \rightarrow B$,

does attribute closure of A include B when $A \rightarrow B$ is excluded from derivation?

if so, eliminate $A \rightarrow B$ and continue

$S \rightarrow N, S \rightarrow A, V \rightarrow T, V \rightarrow Ma, V \rightarrow Mo, Mo \rightarrow Ma, Mo \rightarrow T, V \rightarrow S$

Attribute closure of S without $S \rightarrow N$ is $\{S, A\}$: have to keep $S \rightarrow N$

Attribute closure of S without $S \rightarrow A$ is $\{S, N\}$: have to keep $S \rightarrow A$

Attribute closure of V without $V \rightarrow T$ is still $\{V, Mo, Ma, T, S, N, A\}$: can remove $V \rightarrow T$

Attribute closure of V without $V \rightarrow Ma$ (and without $V \rightarrow T$) is still $\{V, Mo, Ma, T, S, N, A\}$:
can remove $V \rightarrow Ma$

Confirm that we would have to keep $V \rightarrow Mo, Mo \rightarrow Ma, Mo \rightarrow T, V \rightarrow S$

Consider FDs of form $X \rightarrow B$, where X is a set of attributes and B is a single attribute. For example, we might assert $N, A \rightarrow S$ and $V, Mo \rightarrow Ma$ in addition to the other FDs asserted previously (but excluding $V \rightarrow Ma$ and $Mo \rightarrow Ma$).

$S \rightarrow N, S \rightarrow A, V \rightarrow T, V \rightarrow Ma, V \rightarrow Mo, Mo \rightarrow Ma, Mo \rightarrow T, V \rightarrow S,$
 $N, A \rightarrow S, V, Mo \rightarrow Ma$

Complete Alg. For determining minimal FD set

For each FD, $X \rightarrow B$, with non-unit LHS, X

if any proper subset of X , Y , determines remainder of X ($X - Y$), then replace X by Y (replace $X \rightarrow B$ by $Y \rightarrow B$)

$V, Mo \rightarrow Ma$ replaced by $V \rightarrow Ma$

Give an ordering of the FDs (different orderings may result in different minimal FD sets)

For each FD, $Z \rightarrow B$,

does *generalized* attribute closure of Z include B when $Z \rightarrow B$ is excluded from derivation? (e.g., generalized attribute closure of $\{N, A\}$ is $\{N, A, S\}$)

if so, eliminate $Z \rightarrow B$ from set of FDs and continue (*$V \rightarrow T$ eliminated; $V \rightarrow Ma$ not*)

$S \rightarrow N, S \rightarrow A, V \rightarrow Mo, Mo \rightarrow Ma, Mo \rightarrow T, V \rightarrow S$ (or $S \rightarrow N, A; Mo \rightarrow Ma, T; V \rightarrow Mo, S$)

<u>N</u>	<u>A</u>	<u>S</u>	<u>V</u>	<u>T</u>	<u>Ma</u>	<u>Mo</u>
Fred	Nashville	123	987	Truck	Ford	Ranger
Sri	NewYork	234	876	Car	Toyota	Camry
Gabriel	Nashville	345	765	MotorCy	Harley	Hog
Fred	Nashville	123	654	Car	VW	Bug

Show a decomposition of the mega relation into BCNF relations using the FDs above

Give the key(s) of the relation above, as dictated by the FDs.

V is NOT on the right-hand side (RHS) of any FD. Thus, V must be part of any key (i.e., the only way to infer V is to be given V).

$$V \rightarrow Mo \quad Mo \rightarrow Ma$$

Is V alone a key? Yes. Attribute closure of V is $\{V\} \rightarrow \{V, Mo\} \rightarrow \{V, Mo, Ma\}$

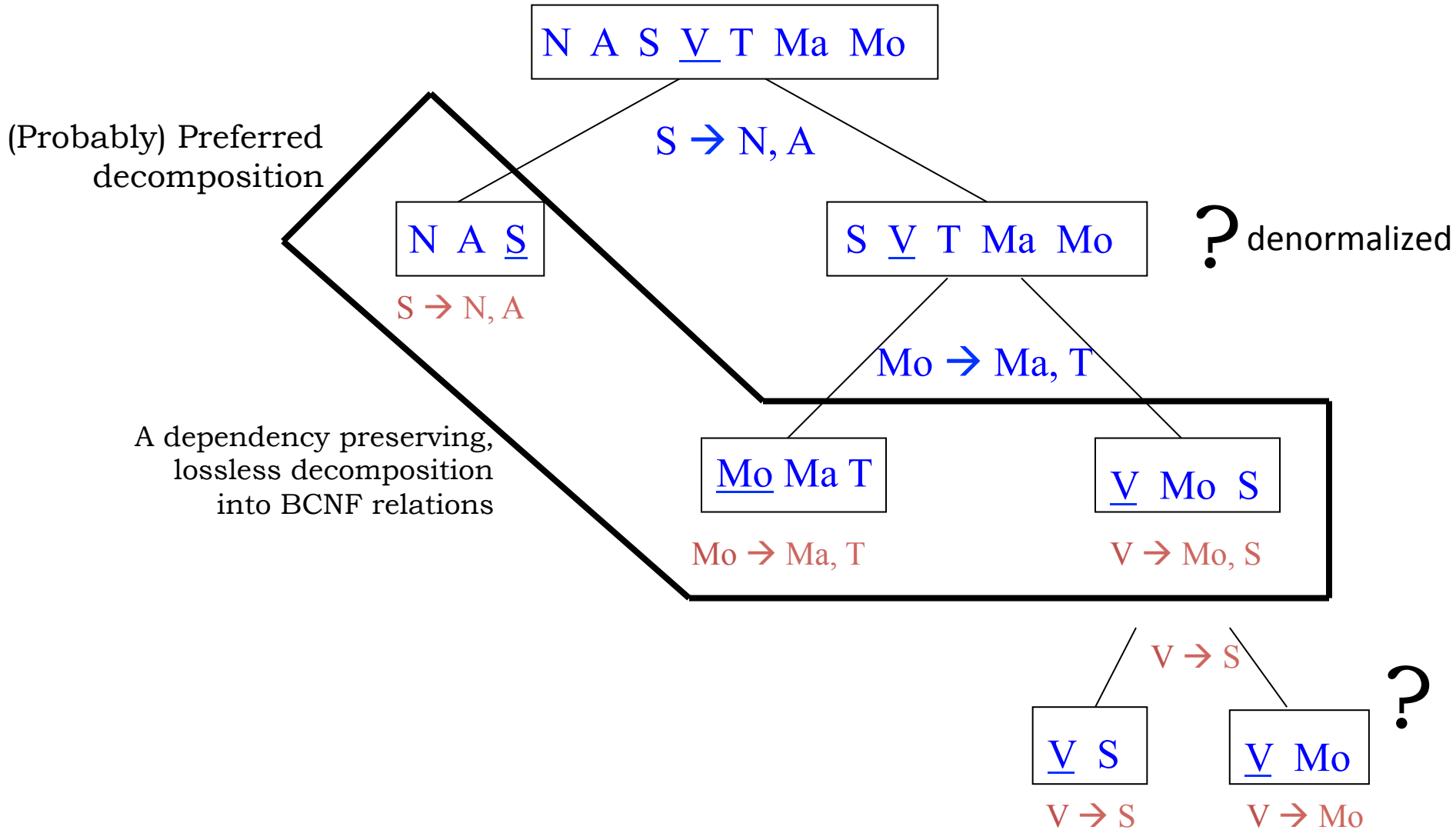
$$Mo \rightarrow T \quad V \rightarrow S \\ \rightarrow \{V, Mo, Ma, T\} \rightarrow \{V, Mo, Ma, T, S\}$$

Are there any other (minimal) keys?

$$S \rightarrow N \quad S \rightarrow A \\ \rightarrow \{V, Mo, Ma, T, S, N\} \rightarrow \underline{\{V, Mo, Ma, T, S, N, A\}}$$

Importance of lossless decomposition: obvious, don't want to lose information

Importance of dependency preservation: each FD constraint can be checked by looking within a single table/relation (i.e., efficiency)



Write CREATE TABLE statements for each of the three relations in the preferred decomposition

<u>N</u>	A	S	V	T	Ma	Mo
Fred	Nashville	123	987	Truck	Ford	Ranger
Sri	NewYork	234	876	Car	Toyota	Camry
Gabriel	Nashville	345	765	MotorCy	Harley	Hog
Fred	Nashville	123	654	Car	VW	Bug

<u>N</u>	A	S
Fred	Nashville	123
Sri	NewYork	234
Gabriel	Nashville	345

<u>T</u>	Ma	Mo
Truck	Ford	Ranger
Car	Toyota	Camry
MotorCy	Harley	Hog
Car	VW	Bug

<u>S</u>	V	Mo
123	987	Ranger
234	876	Camry
345	765	Hog
123	654	Bug

```
CREATE TABLE Person (
    Name VARCHAR(60) NOT NULL,
    Address VARCHAR(120) NOT NULL,
    SSN INTEGER PRIMARY KEY
);
```

Give a UML diagram that is consistent with these Table definitions.

```
CREATE TABLE Description (
    Model CHAR(20) PRIMARY KEY,
    Manufacturer CHAR(20) NOT NULL,
    Type CHAR(10)
);
```

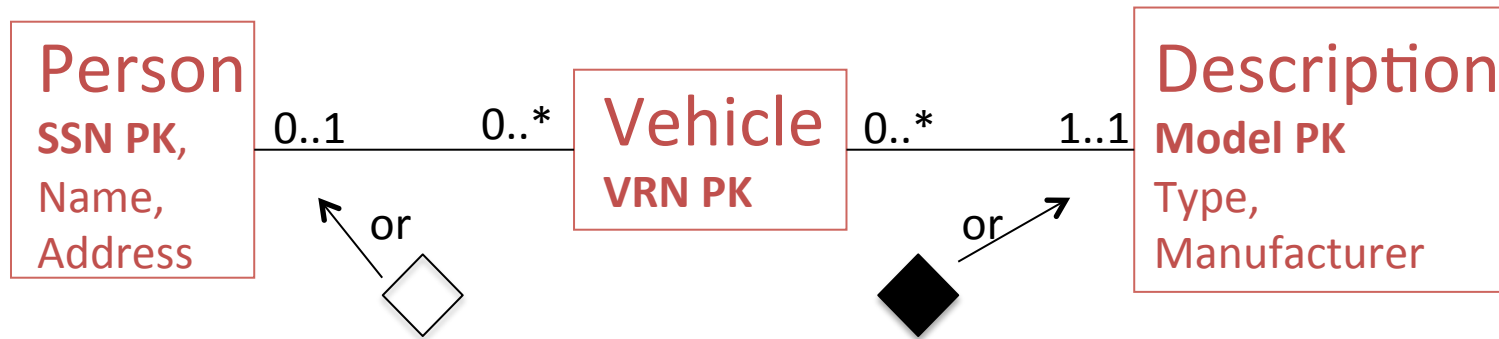


S → N, A

Mo → Ma, T

V → Mo, S

```
CREATE TABLE Vehicle (
    SSN INTEGER, /* NOT NULL? */
    VRN INTEGER PRIMARY KEY,
    Model CHAR(10) NOT NULL,
    FOREIGN KEY (SSN) REFERENCES Person ON DELETE NO ACTION ON UPDATE CASCADE
    FOREIGN KEY (Model) REFERENCES Description ON DELETE NO ACTION ON UPDATE CASCADE
);
```



```

CREATE TABLE Person (
  Name VARCHAR(60) NOT NULL,
  Address VARCHAR(120) NOT NULL,
  SSN INTEGER PRIMARY KEY);
  
```

```

CREATE TABLE Description (
  Model CHAR(20) PRIMARY KEY,
  Manufacturer CHAR(20) NOT NULL,
  Type CHAR(10));
  
```

```

CREATE TABLE Vehicle (
  SSN INTEGER, /* NOT NULL? */
  VRN INTEGER PRIMARY KEY,
  Model CHAR(10) NOT NULL,
  FOREIGN KEY (SSN) REFERENCES Person ON DELETE NO ACTION ON UPDATE CASCADE
  FOREIGN KEY (Model) REFERENCES Description ON DELETE NO ACTION ON UPDATE CASCADE);
  
```