

## **HW#1 Solution**

10.17 Suppose we have the following requirements for a university database that is used to keep track of students transcripts:

(a) The university keeps track of each student's name (SNAME), student number (SNUM), social security number (SSSN), current address (SCADDR) and phone (SCPHONE), permanent address (SPADDR) and phone (SPPHONE), birthdate (BDATE), sex (SEX), class (CLASS) (freshman, sophomore, ..., graduate), major department (MAJORDEPTCODE), minor department (MINORDEPTCODE) (if any), and degree program (PROG) (B.A., B.S., ..., Ph.D.). Both ssn and student number have unique values for each student.

(b) Each department is described by a name (DEPTNAME), department code (DEPTCODE), office number (DEPTOFFICE), office phone (DEPTPHONE), and college (DEPTCOLLEGE). Both name and code have unique values for each department.

(c) Each course has a course name (CNAME), description (CDESC), code number (CNUM), number of semester hours (CREDIT), level (LEVEL), and offering department (CDEPT). The value of code number is unique for each course.

(d) Each section has an instructor (INSTUCTORNAME), semester (SEMESTER), year (YEAR), course (SECCOURSE), and section number (SECNUM). Section numbers distinguish different sections of the same course that are taught during the same semester/year; its values are 1, 2, 3, ...; up to the number of sections taught during each semester.

(e) A grade record refers to a student (Ssn), refers to a particular section, and grade (GRADE).

Design an relational database schema for this database application. First show all the functional dependencies that should hold among the attributes. Then, design relation schemas for the database that are each in 3NF or BCNF. Specify the key attributes of each relation. Note any unspecified requirements, and make appropriate assumptions to make the specification complete.

**Answer:**

From the above description, we can presume that the following functional dependencies hold on the attributes:

FD1: {SSSN} -> {SNAME, SNUM, SCADDR, SCPHONE, SPADDR, SPPHONE, BDATE, SEX, CLASS, MAJOR, MINOR, PROG}

FD2: {SNUM} -> {SNAME, SSSN, SCADDR, SCPHONE, SPADDR, SPPHONE, BDATE, SEX, CLASS, MAJOR, MINOR, PROG}

FD3: {DEPTNAME} -> {DEPTCODE, DEPTOFFICE, DEPTPHONE, DEPTCOLLEGE}

FD4: {DEPTCODE} -> {DEPTNAME, DEPTOFFICE, DEPTPHONE, DEPTCOLLEGE}

FD5: {CNUM} -> {CNAME, CDESC, CREDIT, LEVEL, CDEPT}

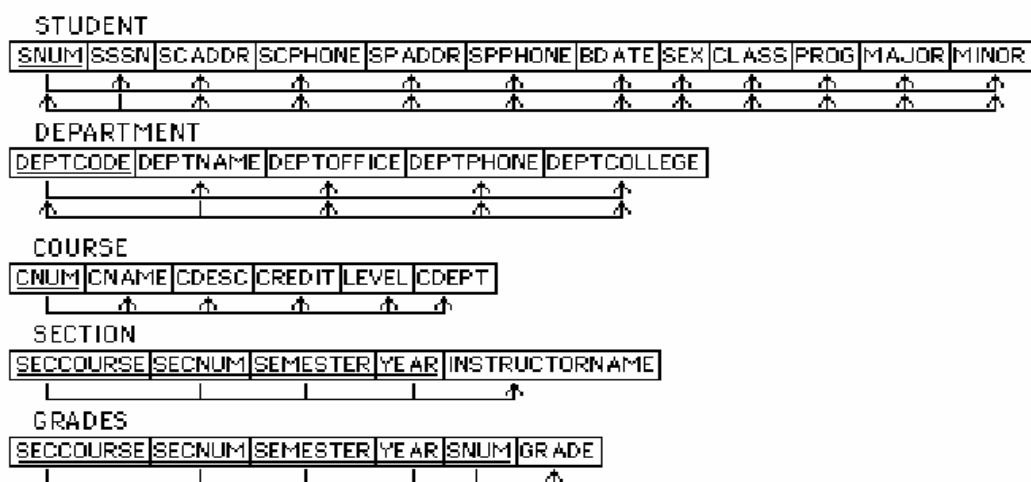
FD6: {SECCOURSE, SEMESTER, YEAR, SECNUM} -> {INSTRUCTORNAME}

FD7: {SECCOURSE, SEMESTER, YEAR, SECNUM, SSSN} -> {GRADE}

These are the basic FDs that we can define from the given requirements; using inference rules IR1 to IR3, we can deduce many others. FD1 and FD2 refer to student attributes; we can define a relation STUDENT and choose either SSSN or SNUM as its primary key.

Similarly, FD3 and FD4 refer to department attributes, with either DEPTNAME or DEPTCODE as primary key. FD5 defines COURSE attributes, and FD6 SECTION attributes.

Finally, FD7 defines GRADES attributes. We can create one relation for each of STUDENT, DEPARTMENT, COURSE, SECTION, and GRADES as shown below, where the primary keys are underlined. The COURSE, SECTION, and GRADES relations are in 3NF and BCNF if no other dependencies exist. The STUDENT and DEPARTMENT relations are in 3NF and BCNF according to the general definition given in Sections 14.4 and 14.5, but not according to the definitions of Section 14.3 since both relations have secondary keys.



The foreign keys will be as follows:

STUDENT.MAJOR -> DEPARTMENT.DEPTCODE

STUDENT.MINOR -> DEPARTMENT.DEPTCODE

COURSE.CDEPT -> DEPARTMENT.DEPTCODE

SECTION.SECCOURSE -> COURSE.CNUM

GRADES.(SECCOURSE, SEMESTER, YEAR, SECNUM) ->

SECTION.(SECCOURSE, SEMESTER, YEAR, SECNUM)

GRADES.SNUM  $\rightarrow$  STUDENT.SNUM

Note: We arbitrarily chose SNUM over SSSN for primary key of STUDENT, and DEPTCODE over DEPTNAME for primary key of DEPARTMENT.

10.18 Prove or disprove the following inference rules for functional dependencies. A proof can be made either by a proof argument or by using inference rules IR1 through IR3. A disproof should be done by demonstrating a relation instance that satisfies the conditions and functional dependencies in the left hand side of the inference rule but do not

satisfy the conditions or dependencies in the right hand side.

(a)  $\{W \rightarrow Y, X \rightarrow Z\} \models \{WX \rightarrow Y\}$

(b)  $\{X \rightarrow Y\}$  and  $Z \subseteq Y \models \{X \rightarrow Z\}$

(c)  $\{X \rightarrow Y, X \rightarrow W, WY \rightarrow Z\} \models \{X \rightarrow Z\}$

(d)  $\{XY \rightarrow Z, Y \rightarrow W\} \models \{XW \rightarrow Z\}$

(e)  $\{X \rightarrow Z, Y \rightarrow Z\} \models \{X \rightarrow Y\}$

(f)  $\{X \rightarrow Y, XY \rightarrow Z\} \models \{X \rightarrow Z\}$

**Answer:**

(a) Proof:

(1)  $W \rightarrow Y$  (given)

(2)  $X \rightarrow Z$  (given)

(3)  $WX \rightarrow YX$  (using IR2 (augmentation) to augment (1) with X)

(4)  $YX \rightarrow Y$  (using IR1 (reflexivity), knowing that  $Y \subseteq YX$ )

(5)  $WX \rightarrow Y$  (using IR3 (transitivity) on (3) and (4))

(b) Proof:

(1)  $X \rightarrow Y$  (given)

(2)  $Y \rightarrow Z$  (using IR1 (reflexivity), given that  $Z \subseteq Y$ )

(3)  $X \rightarrow Z$  (using IR3 (transitivity) on (1) and (2))

(c) Proof:

(1)  $X \rightarrow Y$  (given)

(2)  $X \rightarrow W$  (given)

(3)  $WY \rightarrow Z$  (given)

(4)  $X \rightarrow XY$  (using IR2 (augmentation) to augment (1) with  $X$ )

(5)  $XY \rightarrow WY$  (using IR2 (augmentation) to augment (2) with  $Y$ )

(6)  $X \rightarrow WY$  (using IR3 (transitivity) on (4) and (5))

(7)  $X \rightarrow Z$  (using IR3 (transitivity) on (6) and (3))

(d) Disproof:  $X \ Y \ Z \ W$

$t_1 = x_1 \ y_1 \ z_1 \ w_1$

$t_2 = x_1 \ y_2 \ z_2 \ w_1$

The above two tuples satisfy  $XY \rightarrow Z$  and  $Y \rightarrow W$  but do not satisfy  $XW \rightarrow Z$

(e) Disproof:  $X \ Y \ Z$

$t_1 = x_1 \ y_1 \ z_1$

$t_2 = x_1 \ y_2 \ z_1$

The above two tuples satisfy  $X \rightarrow Z$  and  $Y \rightarrow Z$  but do not satisfy  $X \rightarrow Y$

(f) Proof:

(1)  $X \rightarrow Y$  (given)

(2)  $XY \rightarrow Z$  (given)

(3)  $X \rightarrow XY$  (using IR2 (augmentation) to augment (1) with  $X$ )

(4)  $X \rightarrow Z$  (using IR3 (transitivity) on (3) and (2))

10.26 Consider the universal relation  $R = \{A, B, C, D, E, F, G, H, I\}$  and the set of functional dependencies  $F = \{ \{A, B\} \rightarrow \{C\}, \{A\} \rightarrow \{D, E\}, \{B\} \rightarrow \{F\}, \{F\} \rightarrow \{G, H\}, \{D\} \rightarrow \{I, J\} \}$ . What is the key for R? Decompose R into 2NF, then 3NF relations.

**Answer:**

A minimal set of attributes whose closure includes all the attributes in R is a key.

(One can also apply algorithm 15.4a (see chapter 15 in the textbook)). Since the closure of  $\{A, B\}$ ,  $\{A, B\}^+ = R$ , one key of R is  $\{A, B\}$  (in this case, it is the only key).

To normalize R intuitively into 2NF then 3NF, we take the following steps

(alternatively, we can apply the algorithms discussed in Chapter 15):

First, identify partial dependencies that violate 2NF. These are attributes that are functionally dependent on either parts of the key,  $\{A\}$  or  $\{B\}$ , alone. We can calculate the closures  $\{A\}^+$  and  $\{B\}^+$  to determine partially dependent attributes:

$\{A\}^+ = \{A, D, E, I, J\}$ . Hence  $\{A\} \rightarrow \{D, E, I, J\}$  ( $\{A\} \rightarrow \{A\}$  is a trivial dependency)

$\{B\}^+ = \{B, F, G, H\}$ , hence  $\{B\} \rightarrow \{F, G, H\}$  ( $\{B\} \rightarrow \{B\}$  is a trivial dependency)

To normalize into 2NF, we remove the attributes that are functionally dependent on part of the key (A or B) from R and place them in separate relations R1 and R2, along with the part of the key they depend on (A or B), which are copied into each of these relations but also remains in the original relation, which we call R3 below:

$R1 = \{A, D, E, I, J\}$ ,  $R2 = \{B, F, G, H\}$ ,  $R3 = \{A, B, C\}$

The new keys for R1, R2, R3 are underlined. Next, we look for transitive dependencies in R1, R2, R3. The relation R1 has the transitive dependency  $\{A\} \rightarrow \{D\} \rightarrow \{I, J\}$ , so we remove the transitively dependent attributes  $\{I, J\}$  from R1 into a relation R11 and copy the attribute D they are dependent on into R11. The remaining attributes are kept in a relation R12. Hence, R1 is decomposed into R11 and R12 as follows:

$R11 = \{D, I, J\}$ ,  $R12 = \{A, D, E\}$

The relation R2 is similarly decomposed into R21 and R22 based on the transitive dependency  $\{B\} \rightarrow \{F\} \rightarrow \{G, H\}$ :

$R2 = \{F, G, H\}$ ,  $R2 = \{B, F\}$

The final set of relations in 3NF are  $\{R11, R12, R21, R22, R3\}$

10.29 Given relation R(A,B,C,D,E) with dependencies

$AB \rightarrow C$

$CD \rightarrow E$

$DE \rightarrow B$

is AB a candidate key of this relation?

If not, is ABD? Explain your answer.

**Answers:**

No,  $AB^+ = \{A, B, C\}$ , a proper subset of  $\{A, B, C, D, E\}$

Yes,  $ABD^+ = \{A, B, C, D, E\}$