EQUIVALENT AND
REDUNDANT FUNCTIONAL DEPENDENCIES

Assume we are given a schema R and a set of fds on Attrb(R).

Closure of a Set of Attributes Based on A Set of FD's:

- For each non-empty $X \subset$ Attrb(R), $X^+ = \{Y \in$ Attrb(R): $X \rightarrow Y\}$; X^+ is called the *closure* of X.

  Here, $X \rightarrow Y$ may be in F or it may be derivable from F via the transitive and other rules on fd's.

Properties of Closure Operation:

- *Upwardness:* $X \subseteq X^+$ (because of trivial fd's) and $\varnothing^+ = \varnothing$.
- *Monotonicity:* If $X \subseteq X'$, then $X^+ \subseteq (X')^+$.
- *Idempotent Property:* $(X^+)^+ = X^+$.

Example. Let $F = \{A \rightarrow B,\ A \rightarrow C,\ C \rightarrow A\}$.

- We write a set of attributes by listing them together, without the enclosing in '{' and '}' and the separating commas.

$$A^+ = ABC = (AB)^+ = (AC)^+ = (ABC)^+$$
$$B^+ = B$$
$$C^+ = ABC$$

Question:

•? What is $(BC)^+$ and $[(BC)^+]^+$ for the F above?

•? Show the closures X^+ for $F = \{A \rightarrow B,\ B \rightarrow C,\ C \rightarrow A\}$. How do you compare these closures X^+ with those in Example above? Why does this happen?

CLOSURE OF A SET OF fds

Closure F^+ of a set F of fds:

- F^+ consists of all fds of the form $X \to X^+$, $X \neq \emptyset$. Put another way, F^+ is the set F together with all other non-trivial fd's that can be obtained by transitive and other properties of fd's.
- If $F_1 \subseteq F_2$, then X^+ w.r.t F_1 is a subset of X^+ w.r.t F_2 and thus $F_1^+ \subseteq F_2^+$. Also, $F^+ = (F^+)^+$ and $\emptyset^+ = \emptyset$.
- For each X, X^+ is the same w.r.t F and F^+.

Question: If $F = \emptyset$, then what is X^+?

Example.

- Consider $F_1 = \{A \to B,\ A \to C,\ C \to A\} \subseteq \{A \to B,\ B \to C,\ A \to C,\ C \to A\} = F_2$.

  Then, $F_1^+ = \{A \to ABC,\ B \to B,\ C \to ABC\}$. (What is F_2^+?)
- In both F_1 and F_1^+, $B^+ = B$; in both F_2 and F_2^+, $B^+ = ABC$.

Equivalence of F_1 and F_2 ($F_1 \cong F_2$):

- F_1 is *equivalent* to F_2 if $F_1^+ = F_2^+$, which is the same as $F_1 \subseteq F_2^+$ and $F_2 \subseteq F_1^+$. (Why?)
- To test $F_1 \subseteq F_2^+$, we need to verify that each fd $X \to Y \in F_1$ is in F_2^+, i.e., $Y \subseteq X^+$ w.r.t F_2. (No need to compute F_2^+.)

Question:

•? Show that $F_1 = \{A \to B,\ B \to C,\ C \to A\}$ and $F_2 = \{A \to C,\ C \to B,\ B \to A\}$ are equivalent.

•? Argue that $F_1 \cong F_2$ if and only if for each $X \subseteq$ Attrb(R) we have X^+ is the same w.r.t F_1 and F_2.

REDUNDANT fds

Redundancy:

- An fd f: $X \rightarrow Y \in F$ is *redundant* if $F \cong F - \{f\}$, i.e $F^+ = (F - \{f\})^+$.
- This is the same as saying $X \rightarrow Y$ can be obtained from the other fds in F. This, in turn, is the same as saying $Y \subseteq X^+$ computed using $F - \{f\}$.

Example.

- $A \rightarrow C$ is redundant in $\{A \rightarrow B,\ B \rightarrow C,\ A \rightarrow C\}$ because $A^+ = ABC$ w.r.t $\{A \rightarrow B,\ B \rightarrow C\}$.

Reduced F:

- A set of fd's F is *reduced* if it contains no redundant fd.

Example.

- The set $F = \{A \rightarrow B,\ B \rightarrow C,\ C \rightarrow A,\ A \rightarrow C,\ C \rightarrow B,\ B \rightarrow A\}$ can be reduced in many ways; shown below are some.

1. $\{A \rightarrow B,\ B \rightarrow C,\ C \rightarrow A\}$
2. $\{A \rightarrow C,\ C \rightarrow B,\ B \rightarrow A\}$
3. $\{A \rightarrow C,\ C \rightarrow A,\ C \rightarrow B,\ B \rightarrow C\}$

Question:

•? Show all possible reduced form of F in Example above.

BERNSTEIN'S ALGORITHM

Algorithm Form-3NF-Decomposition:

Input: A schema R and a set F of fds on Attrb(R) = $\{A_i: 1 \le i \le n\}$.

Output: A set of 3NF relations that form a loss-less, fd-preserving decomposition.

1. Reduce F by eliminating all redundant fd's.
2. Group together all fd's which have the same left-hand side. That is, if $X \to A_1$, $X \to A_2$, $\cdots$, $X \to A_k$ are the fd's with the l.h.s X, then replace them by the single fd $X \to A_1A_2 \cdots A_k$,
3. For each fd $X \to Y$, form the relation schema (XY) with the attributes $X \cup Y$ in the decomposition (representing $\Pi_{XY}(R)$).
4. If $X'Y' \subset XY$, remove schema $(X'Y')$ from the decomposition.
5. If no schema in the decomposition contains a key of R, then add schema (K), where K is a key, to the decomposition.

Example. It shows the need for steps (2) and (4).

- Attrb(R) = $\{A, B, C\}$ and $F = \{A \to B, A \to C, C \to A, C \to B\}$.

  Step 1. $\{A \to B, A \to C, C \to A\}$
  Step 2. $\{A \to BC, C \to A\}$
  Step 3. (ABC), (CA)
  Step 4. (ABC)
  Step 5. (ABC); the key A is contained in ABC.

- Without steps (2) and (4), we get the decomposition $(ABC) = (AB) \otimes (AC)$, which is not preferred (although it is lossless, fd-preserving, and each of (AB) and (AC) is in 3NF), because it causes unnecessary duplication of values for the attribute A. The relation (ABC) itself is in 3NF.

ILLUSTRATION OF (ABC) = (AB) ⊗ (AC)

A	B	C		A	B		A	C
a1	b1	c1		a1	b1		a1	c1
a2	b1	c2	=	a2	b1	⊗	a2	c2
a3	b3	c3		a3	b3		a3	c3

It satisfies $\{A \rightarrow B,\ A \rightarrow C,\ C \rightarrow A\}$.

ANOTHER ILLUSTRATION OF BERNSTEIN'S ALGORITHM

Example. It shows the need for step (5).

- Let Attrb(R) = $\{A, B, C, D\}$ and $F = \{A \to C, B \to D, AB \to CD\}$.

Step 1.	$\{A \to C, B \to D\}$
Step 2.	$\{A \to C, B \to D\}$
Step 3.	$(AC), (BD)$
Step 4.	$(AC), (BD)$
Step 5.	$(AC), (BD), (AB)$

A	B	C	D		A	C		B	D
a1	b1	c1	d1	≠	a1	c1	⊗	b1	d1
a1	b2	c1	d2		a2	c2		b2	d2
a2	b2	c2	d2						

but it is easy to see that $(ABCD) = (AC)\otimes(BD)\otimes(AB)$.

Question:

•? Show $(AC)\otimes(BD)$ in the above example, and verify that $(ABCD) = (AC)\otimes(BD)\otimes(AB)$.

•? What are the possible sets of fds F such that $(ABC) = (AB)\xi(BC)$ is a Bernstein's decomposition?

ANOTHER ILLUSTRATION

Example.

- Consider a course-schedule with the attributes course(C), student(S), teacher(T), class-period-time(P), room(R), and day(D). Here, C includes the course-section, if any.
- The fds and their meaning are:

(1) $C \to P$:	A course meets at the same period each time.
(2) $C \to R$:	A course meets at the same room each time.
(3) $C \to T$:	No joint teaching of any course.
(4) $TPD \to C$:	A teacher cannot teach more than one course at a period of any day.
(5) $TPD \to R$:	A teacher cannot be in more than one room at a period of any day.
(6) $SPD \to C$:	A student cannot take more than one course at a period of any day.
(7) $SPD \to R$:	A student cannot be in more than one room at a period of any day.

It is easy to see that (4) and (2) implies (5), and (6) and (2) implies (7). A reduced set of fd's is

(1)	C → P	(2)	C → R
(3)	C → T	(4)	TPD → C
(5)	SPD → C		

There are two keys: SDP, SDC. The Bernstein's fd-preserving, loss-less, 3NF decomposition is

(CDPRST) = (CPRT) ⊗ (TPDC) ⊗ (SPDC)
= (CP) ⊗ (CR) ⊗ (CT) ⊗ (TPDC) ⊗ (SPDC)
= (CR) ⊗ (TPDC) ⊗ (SPDC). (Bernstein's method may not give the simplest 3NF decomposition.)

Some Impportant Comments:

1. The relation (CP) is deleted because the attribute $\{C, P\}$ (in short, CP) appear in other relations. Actually (CP) can be omitted only by assuming that every CP-tuple has at least some student or some instructor associated with it.

   This would not be the case if (CP) represents a tentative schedule before the students and instructors are assigned.

2. Similar remarks hold for (CT).

3. The relation ($SPDC$) is what the students need to know, and the relation ($TPDC$) is what the teachers need to know. Note that there is significant amount of duplication of CPD-information in between ($SPDC$) and ($TPDC$).

4. We should not merge (TPDC) and (SPDC) into one relation (STPDC) because the latter is not 3NF under the fd's $\{C \rightarrow PT, SPD \rightarrow C, TPD \rightarrow C\}$ which give the keys SDC and SDP. The relation (TPDC) with the fd's $\{C \rightarrow PT, TPD \rightarrow C\}$ is 3NF as is (SPDC) with $\{C \rightarrow P, SPD \rightarrow C\}$.

5. The relation (CR) is useful to both teachers and students.