3NF - Third Normal Form

Definition 1:
A relation schema is in 3NF if
(1) it is in 2NF and
(2) each non-key attribute must <u>not</u> be fully functionally dependent on another non-key attribute (there must be no transitive dependency of a non-key attribute on the PK).

Definition 2:
A relation schema R is in 3NF if, whenever a function dependency $X \rightarrow A$ holds in R, either

(a) X is a superkey of R, or
(b) A is a prime attribute of R.

A superkey of a relation schema R = {A1, A2, ..., An} is a set of attributes $S \subseteq R$ with the propertity that no tuples t1 and t2 in any legal state r of R will have t1[S] = t2[S].

An attribute is called a prime attribute if it is a member of any key.

Theorem Definition 1 and Definition 2 are equivalent.

Proof.

1. If R is in 3NF according to Definition 2, R must be in 3NF according to Definition1.

Recall that if R is in 3NF according to Definition 1, then the following two conditions must be satisfied.

i) For R, we don't have any transitive function dependency between a non-key attribute and a key through some other non-key attribute.
ii) For R, we don't have any partial function dependency between a non-key attribute and a key.

Assume that R does not satisfy (i). Then, we must have a transitive FD: $X \rightarrow A$, $A \rightarrow B$, where X is a key, A and B are non-key attributes. But according to Definition 2, R does not have such kind of FDs. (That is, A must be a prime attribute or a super key.) Contradiction. So R must be satisfy (i).
Assume that R does not satisfy (ii). Then, we must have a subset S of a key X ($S \subset X$) such that there exists a non-key attribute A with $S \rightarrow A$. However, according to Definition 2, S must be a super key. Contradiction. So R must satisfy (ii).

Therefore, R is in 3NF according to Definition 1.

2. If R is in 3NF according to Definition 1, R must be in 3NF according to Definition 2.

Assume that R is not in 3NF according to Definition 2. Then, we must have a FD: X $\rightarrow$ A such that X is not a super key and A is not a prime attribute. Consider the a key X' of R. We must have

$$X' \rightarrow X \rightarrow A.$$

It is a transitive FD between the non-key attribute A and the key X' through X. If X is a non-key attribute, then R is not in 3NF according to Definition 1. Contradiction. If X appears in X', we have a partial FD. So R is not in 2NF, contradicting to Definition 1.