Normalization Levels and Normalizing a Relation

Tom Kelliher, CS 417

- Superkeys, candidate keys, and primary keys – See review on pg. 265 of the textbook.

- Functional dependence within a relation — $A \to B$. A and B are sets of attributes. We say "A determines B." A is the *determinant* and B is the *dependent*.

- First Normal Form (1NF) — A relation is in first normal form if, and only if, every attribute is single-valued for each tuple.

- Second Normal Form (2NF) — A relation is in second normal form if, and only if, it is in first normal form and all its non-key attributes are *fully functionally dependent* on the key.

  Definition: In a relation R, attribute A of R if *fully functionally dependent* on an attribute or set of attributes X of R if A is functionally dependent on X ($X \to A$) but not functionally dependent on any proper subset of X.

- Third Normal Form (3NF) — A relation is in third normal form if, whenever a non-trivial dependency $X \to A$ exists, then either X is a superkey or A is a member of some candidate key.

    "The essence of third normal form is that each non-key attribute is functionally dependent on the entire key, and on no other attribute."

  Put another way, *each attribute must describe the key, the whole key, and nothing but the key*.

- Boyce-Codd Normal Form (BCNF) — A relation is in Boyce-Codd normal form if, whenever a non-trivial dependency $X \to A$ exists, then X is a superkey.

    "Identify all the determinants and verify that they are superkeys."

- Normalizing a relation – See Sections 6.6 and 6.7 of the textbook.