02 Database normalization

Database normalization

© Janusz R. Getta CSCl235/MCS9235/CSCl835 Databases, SCIT, Autumn 2015

02 Database normalization

First normal form (1NF)

Relational schema is in 1NF if all occurrences of rows in the respective relational table contain the same number of fields and include the atomic values only, (there is no repeating fields and groups)

e#	name	car used
950001	Peter	Toyota, PKR234, Ford, WER545
932345	Paul	Honda, RTQ456
960020	Joan	Holden, KLR197, Holden, KLR567

CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015

02 Database normalization

Keys

Superkey

Superkey is a nonempty subset x of relational schema $R=(A_1,\ldots,A_n)$ such that for any two rows t_1 , t_2 in a relational table defined over R $t_1[X] \neq t_2[X]$ If X is a superkey in R than $X \to (A_1,\ldots,A_n)$

Minimal key

Minimal key is a superkey K with an additional property such that removal of any attribute from K will cause K not to be a superkey

© Janusz R. Getta CSCl235/MCS9235/CSCl835 Databases, SCIT, Autumn 2015 3

02 Database normalization

Keys

Primary key

© Janusz R. Getta

Primary key is an arbitrarily selected minimal key

Candidate key

Candidate key is any other minimal key which is not primary

important key def here

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015

02 Database normalization

Functional dependencies and keys

Let $R=(A_1,\ldots,A_n)$ be a relational schema (a header of relational table) and let X,Y be nonempty subsets of R such that $X\cup Y=R$

- (1) If $X \to Y$ is valid in R then X is a key (!!!)
- (2) If x is a key then $x \rightarrow y$ is valid in R

© Janusz R. Getta CSCl235/MCS9235/CSCl835 Databases, SCIT, Autumn 2015

02 Database normalization

Attributes

Prime attribute

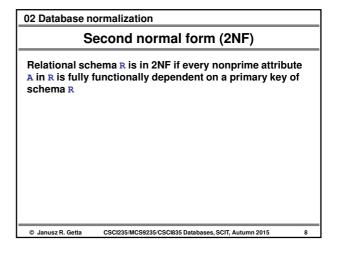
Prime attribute is an attribute from relational schema R which is a member of at least one candidate key in R

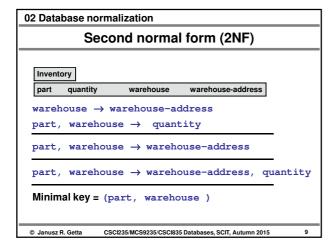
Nonprime attribute

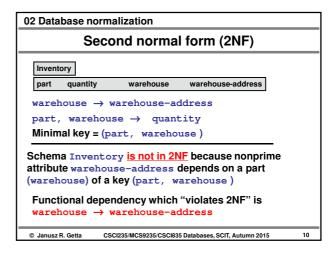
Nonprime attribute is an attribute which is a not prime

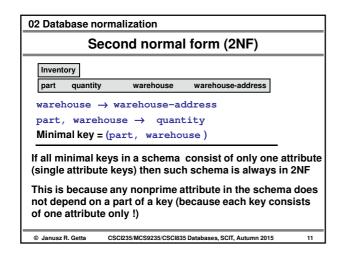
© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 6

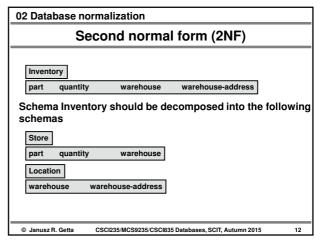
Full and partial functional dependencies Full functional dependency Full functional dependency is a functional dependency X → Y such that removal of any attribute A from X causes that (X-A) → Y Partial functional dependency Partial functional dependency Partial functional dependency is a functional dependency which is not full

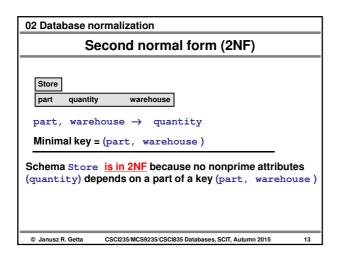


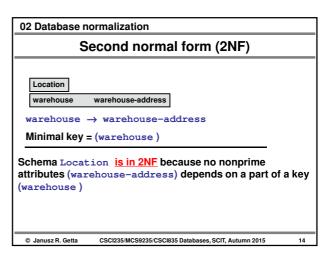


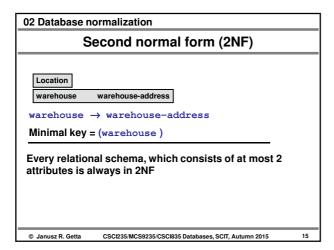


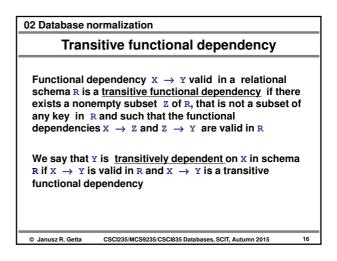




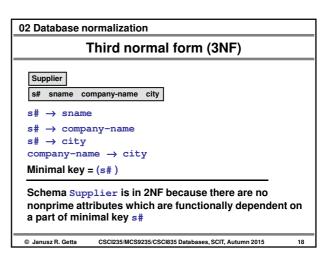


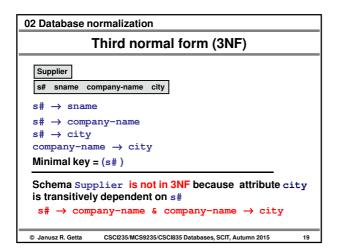


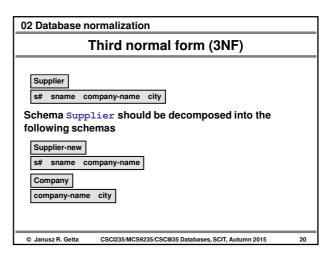


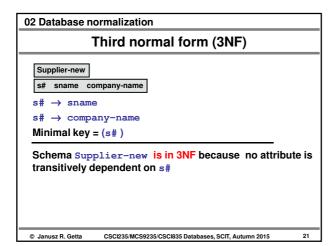


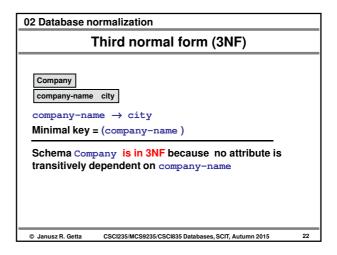
Third normal form (3NF) Relational schema R is in 3NF if it is in 2NF and no nonprime attribute of R is transitively dependent on the primary key © Janusz R. Getta CSCI235/MCS9235/CSCIB35 Databases, SCIT, Autumn 2015 17











Third normal form (3NF)

Company

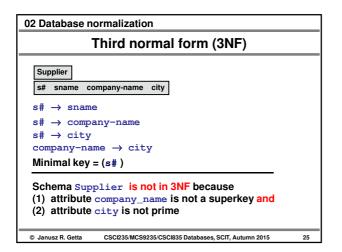
company-name city

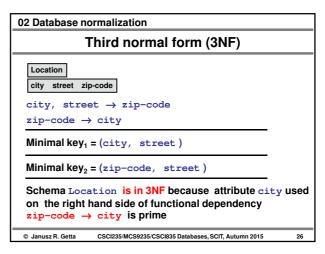
company-name → city

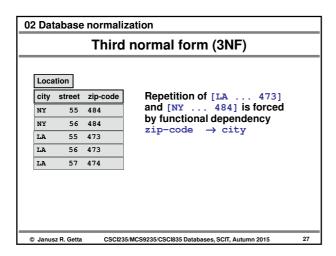
Every relational schema, which consists of at most 2 attributes is always in 3NF

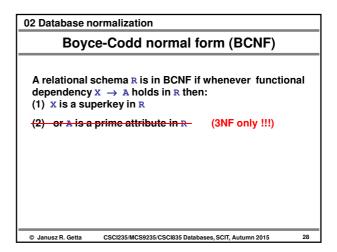
Third normal form (3NF)

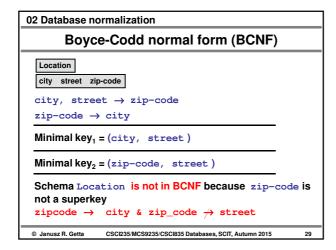
Alternative definition
A relational schema R is in 3NF if whenever a functional dependency X → A is valid in R then either:
(1) X is a superkey in R, or
(2) A is a prime attribute in R

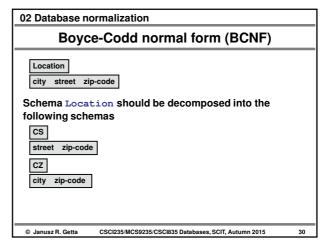


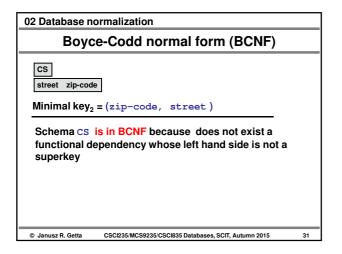


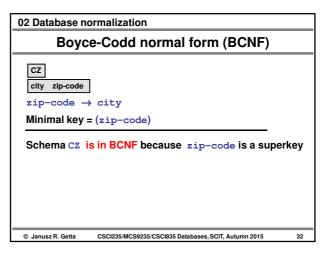


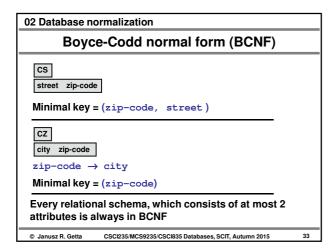


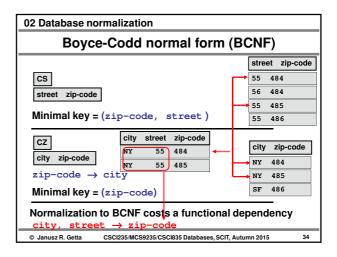




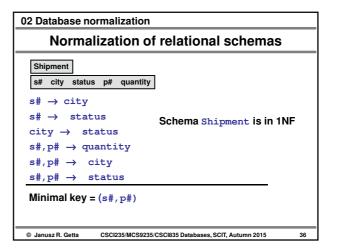


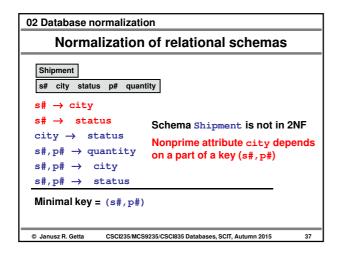


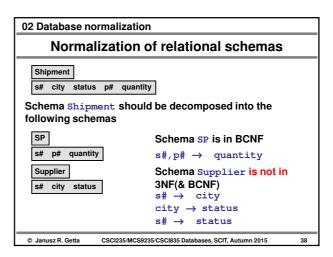


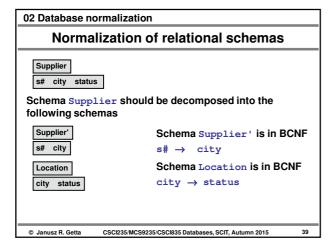


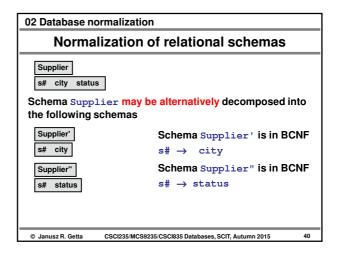
 $\begin{tabular}{ll} \hline \textbf{Normalization of relational schemas} \\ \hline \textbf{R} &= \{ & \textbf{A}_1, & \dots, & \textbf{A}_n \} \\ \hline \textbf{Identify all functional dependencies} \\ \hline \textbf{Use functional dependencies to derive all minimal keys} \\ \hline \textbf{Use functional dependencies and minimal keys to identify the highest normal form satisfied by R} \\ \hline \textbf{Decompose R into the schemas in BCNF (3NF)} \\ \hline \end{tabular}$











References Elmasri R., Navathe S. B., Fundamentals of Database Systems, chapters 10.3,10.4,10.5 R. Ramakrishnan, J. Gehrke Database Management Systems, chapters 19.2, 19.3