Normalisation Theory

→ Normal forms BCNF ⊂ 3NF ⊂ 2NF ⊂ 1NF

* Basic Normal Forms : All attributes are atomic

* BCNF (Boyce - Codd Normal Form)
  - for all $X \to A$ in F^+
    - $A \subseteq X$ (trivial FD) or
    - X is a superkey for R

meaning R is in BCNF if all non trivial FDs are key constraints.

Not a BCNF if $X \to Y$ and X is not a superkey

* 3NF (third normal form)
  - for all $X \to A$ in F^+
    - $A \subseteq X$ or
    - X superkey for R or
    - A part of candidate key (not superkey) for R or (A prime)

minimality of the key is crucial.

Lossless Decomposition:

$r = \pi_X(r) \bowtie \pi_Y(r)$

Decomposition is lossless onto X and Y if $X \cap Y \to Y$ or $X \cap Y \to X$ in F^+

if $X \to Z$ and $X \cap Z = \emptyset$ then $R - Z$ and XZ are lossless.

Projection of a set of FDs

$X = UV$ $U \to V$

$F_X = X$

Proj F onto X is F_X s.t. $U \to V$ in F^+ and $U, V \subset X$

a Decomposition is said to be Dependency preserving (on X, Y)

if $(F_x \cup F_y)^+ = F^+$

Dep. preserving $\not\Rightarrow$ Lossless join

0

$R =$

$F =$

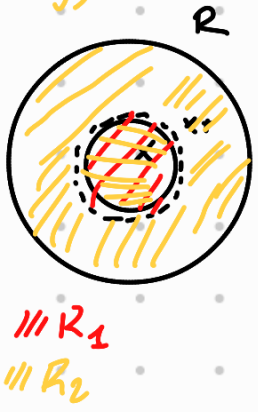
How to Decompose into BCNF:

check if R in BCNF { Yes ok ✓ / No

① pick $X \to Y$ that violates BCNF { $Y \not\subseteq X$ ✗ / X (not a super key)

② → compute X^+

③ → since X^+ is not a super key $X^+ \subset R$.
let $R_1 = X^+$ and $R_2 = X \cup (R - X^+)$



④ → compute FDs for R_1 and R_2

⑤ Repeat ① for R_1 and R_2

Example: CSJDPQV C is a key

JP → C
SD → P
J → S

JP → C ✓

SD → P ✗ → $SD^+ = SDP$

$R_1 = SDP$

$R_2 = SD \cup (CJQV)$
$= CSJDQV$

SDP is BCNF

$J \to S$; $J^+ = JS$ $\begin{cases} R_1 = JS & \checkmark \\ R_2 = CJDQV & \checkmark \end{cases}$

$CSJDPQV \xrightarrow{BCNF} SDP, JS, CJDQV$

BCNF $\not\Rightarrow$ Dependency preserving

↳ How to DO so

to ensure Dependency preservation (in BCNF):

if $X \to Y$ is not preserved add XY ✓

⚠ but XY may violate 3NF $\xrightarrow{\text{mitigation}}$ minimal cover for F

→ 3NF Decomposition

Min. cover G $\begin{cases} G^+ = F^+ \\ G \text{ minimal (not unique)} \end{cases}$

* G helps finding superkeys.

* 3NF lossless join and dependency preserving decomp.

— — — — - - -

if $X \to Y$ in G then add XY to decomposition.