Informal guidelines for Schema Creation

- Semantic understanding: each tuple represents one E/R
- Data redundancy: update, insert, delete anomalies → emp + proj
- NULL values
- Spurious tuples: no duplicate non-key feature

Bottom up: start from individual attributes and combine. Not popular, lot of binary relationships to be captured

Top down: group attributes based on semantic meaning and decompose as needed

Functional dependencies

full, partial, transitive, trivial, non-trivial

$A \to B$: Value of A uniquely determines that of B

A	B
a_1	b_1
a_2	b_1
a_3	b_2

Functional dependency

dept_name → building, budget

Lossy decomposition

Decomposition that results in loss of information; original relation unable to be reconstructed from decomposed relations

NOTE: This includes cases where there is additional unnecessary information on ⋈. Again, you are unable to retrieve the original relation

LOSSLESS JOIN DECOMPOSITION

Decomposition of relation R into R_1, R_2 such that $R_1 \bowtie R_2 = R$

This effectively removes redundancy while preserving original data

$R_1 \cup R_2 = R, \quad R_1 \cap R_2 \neq \phi$

ARMSTRONG'S AXIOMS

F → set of functional dependencies

New dependencies can be inferred by the following axioms:

Reflexivity: if $\beta \subseteq \alpha$, $\alpha \to \beta$

Augmentation: if $\alpha \to \beta$, $\gamma\alpha \to \gamma\beta$ where γ is another set of attributes

Transitivity: if $\alpha \to \beta$ and $\beta \to \gamma$, then $\alpha \to \gamma$

- New FDs created using these on F → logically implied from F
- New FDs generated are finite. Set of F and all logically implied FDs on F: Closure set F^+
- $F \subseteq F^+$

These axioms are sound (only generate FDs that hold) and complete (generate all FDs that hold).

In addition to the above:

Decomposition: $X \to YZ \Rightarrow X \to Y, X \to Z$

Additive / Union: $X \to Y$, $X \to Z \Rightarrow X \to YZ$

Pseudo-transitivity: $X \to Y$, $WY \to Z \Rightarrow WX \to Z$

Derived from the above axioms

Closure of attributes

$(\text{class ID})^+ = \{\text{course\#, instr\_name, credit\_hrs, text, publisher, classroom, capacity, class\_ID}\}$

$(\text{Course\#})^+ = \{\text{course\#, credit\_hrs}\}$

$\{\text{course\#, Instr\_name}\}^+ = \{\text{course\#, instr\_name, credit\_hrs, text, classroom, publisher, capacity}\}$

Closure of FDs

1. reflexive

loop:

2. augmentation
3. transitive

If n attributes in R, $n(FDs) = 2^{2n}$

① $A \to BC$ ② $B \to D$ ③ $C \to E$

$A \to B$, $A \to C$ } Decomposition on ①

④ $A \to D$, ⑤ $A \to E$ } Transitivity on ①-②, ①-③

$A \to DE$ } Union on ④, ⑤

$A \to BC$, $B \to E$, $CD \to EF$

$A \to B$, $A \to C$

$\Rightarrow A \to E$, $AD \to CD$

$\Rightarrow AD \to EF$

$\Rightarrow \underline{AD \to F}$

---

① $SC \to Z$ ② $Z \to C$

③ $SZ \to SC$ (aug from ②)

④ $SC \to SCZ$ (aug from ①)

$\Rightarrow \underline{SZ \to SCZ}$ (transitivity from ③, ④)

---

① $D \to M$ ② $M \to D$

③ $ED \to EM$ (aug ①)

④ $EM \to EDM$ (aug ②, aug ④)

⑤ $ED \to EDM$ (transitivity, ③ and ④)

---

① $AB \to E$ ② $BE \to I$ ③ $E \to G$ ④ $GI \to H$

Show $AB \to GH$

⑤ $AB \to G$ (trans. ①, ③)

⑥ $AB \to BE$ (aug ①)

⑦ $AB \to I$ (trans. ⑥, ②)

⑧ $AB \to GI$ (union ⑤, ⑦)

⑨ $AB \to H$ (trans. ⑧, ④)

⑩ $AB \to GH$ (union ⑤, ⑨)

---

① $AB \to C$ ② $C \to D$ ③ $D \to E$

Find $AB \to E$

$AB \to D$

$AB \to E$

① $A \to B$ ② $A \to C$ ③ $BC \to D$

$AC \to BC$, $A \to AC$

$\to A \to BC$

$\to \underline{A \to D}$

---

① $A \to B$ ② $C \to D$ ③ $BD \to E$

Find $AC \to E$

$AC \to BC$

$AC \to AD$

$AC \to B$, $AC \to D$

$AC \to BD$

$\underline{AC \to E}$

$A \to B$, $BC \to DE$, $AEG \to G$, $AC \to BC$, $AD \to BD$

$\{A\}^+ = \{A, B\}$

$\{AC\}^+ = \{A, C, B, D, E\}$

$\{ADE\}^+ = \{A, D, E, B\}$

$A \to BC$, $CD \to E$, $B \to D$, $E \to A$

$\{B\}^+ = \{B, D\}$

$A \to BC$, $CD \to E$, $D \to AEH$, $E \to C$, $ABH \to BD$, $DH \to BC$

$\{BC\}^+ = \{B, C, D, E, A, H\}$

Minimal Cover of Set of FDs

1. Decompose
2. Remove extraneous
3. Remove redundant: use transitive rule

for each $X \to A$ w/ $|X| > 1$:
  for each Y in X:
    remove Y, NOT FD
    closure (X)
    if $A \in$ closure(X) Y is extraneous

calculate closure for every subset. if any are the same: that's sufficient

for each $X \to A$:
  Remove FD
  closure (X)
  if $A \in$ closure(X) now, redundant

E: $B \to A$, $D \to A$, $AB \to D$

1. Already in canonical form
2. $B \to AB \to D$
 $\Rightarrow$ Replace $AB \to D$ w/ $B \to D$
 $E = \{B \to A, D \to A, B \to D\}$
3. $B \to D$ & $D \to A$
 $\Rightarrow B \to A$ is redundant

Minimal $E = \{D \to A, B \to D\}$

$A \to C$, $AB \to C$, $C \to DI$, $CD \to I$, $EC \to AB$, $EI \to C$

$\Rightarrow$ $A \to C$, $AB \to C$, $C \to D$, $C \to I$, $CD \to I$, $EC \to A$, $EC \to B$, $EI \to C$

② $\{AB\}^+ = \{A, B, C\}$ $\Rightarrow AB \to C$ is redundant

$\{CD\}^+ = \{C, D, I\}$ $\Rightarrow C \to I$ is redundant

③ $A \to C$, $C \to D$, $(CD \to I)$, $EC \to A$, $EC \to B$, $EI \to C$

(i) $\{E\}^+ = \{E\}$, $\{C\}^+ = \{C, D, I\}$ no extraneous

(ii) $\{E\}^+ = \{E\}$, $\{C\}^+$

(i) $\{C\}^+ = \{C, D, I\}$, D is extraneous

Equivalence of Sets of FDs

F = $A \to C$, $AC \to D$, $E \to AD$, $E \to H$

G = $A \to CD$, $E \to AH$, $AC \to CD$

For set F:

$\{A\}^+ = \{A, C, D\}$

$\{AC\}^+ = \{A, C, D\}$

$\{E\}^+ = \{E, A, D, H, C\}$

For set G:

$\{A\}^+ = \{A, C, D\}$

$\{E\}^+ = \{E, A, H, C, D\}$

$\{AC\}^+ = \{A, C, D\}$

① $\{A\}^+, \{E\}^+$: FDs of F determine everything that FDs of G do

 $\Rightarrow F \supseteq G$

② $\{A\}^+, \{E\}^+, \{AC\}^+$: G determines everything F did

 $\Rightarrow G \supseteq F$

③ $F \equiv G$

you look at these closures in the set that will become a subset

F covers G: $F \supseteq G$

Finding keys

Finding a key

1. Set K = all attributes
2. Remove attrs one at a time and check closure
 if still spans: remove attr A

Finding candidate keys

Determine essential attributes: those that are not on RHS of any FD

Case 1: All essential attrs can determine all remaining non essential
 → only candidate key

Case 2: " " " cannot
 → try combinations with non-essential

$C \to F$, $E \to A$, $EC \to D$, $A \to B$

$\{CD\}^+ = \{C, D\}$

$\{EC\}^+ = \{E, C, A, B, D, F\}$ ✓ key

$\{AE\}^+ = \{A, E, B\}$

$\{AC\}^+ = \{A, C, B, F\}$

do not have both C, E (essential)

Essential attrs: $\{C, E\}$

$\{CE\}^+ = \{C, E, A, B, D, F\}$ ✓

only 1 candidate key

C E A B D F

has to be there; may or not be: 2 choices per slot, yes/no

$\Rightarrow 2 \times 2 \times 2 \times 2 = 16$ superkeys