Dipendenza funzionale: dato uno schema $R(T)$ e $X, Y \subseteq T$, una dipendenza funzionale (DF) fra gli attributi di X e Y, è un vincolo su R espresso nella forma $X \to Y$ (X determina Y).

$\forall r$ istanza valida di R.

$\forall t_1, t_2 \in r$. se $t_1[X] = t_2[X]$ allora $t_1[Y] = t_2[Y]$

Ogni dipendenza funzionale $X \to A_1, \dots, A_n$ si può scomporre nelle dipendenze funzionali $X \to A_1, X \to A_2, \dots, X \to A_n$. Una dipendenza del tipo $X \to A$ si dice DF atomica.

Una DF è completa se $X \to Y$ e $\forall W \subseteq T$, non vale $W \to Y$.

Se X è superchiave, allora X determina ogni altro attributo della relazione $X \to T$.

Se X è chiave, allora $X \to T$ è una DF completa.

Assiomi di Armstrong:

- Se $Y \subseteq X$, allora $X \to Y$ (Riflessività R)
- Se $X \to Y$, $Z \subseteq T$, allora $XZ \to YZ$ (Arricchimento A)
- Se $X \to Y$, $Y \to Z$, allora $X \to Z$ (Transitività T)

Derivabilità: sia F un insieme di DF, diremo che $X \to Y$ sia derivabile da F ($F \vdash X \to Y$) se $X \to Y$ può essere inferita da F usando gli assiomi di Armstrong. Valgono anche le seguenti regole:

- $\{X \to Y, X \to Z\} \vdash X \to YZ$ (Unione U)
- $Z \subseteq Y$ $\{X \to Y\} \vdash X \to Z$ (Decomposizione D)

Chiusura: dato un insieme F di DF, la chiusura di F, denotata con F^+, è $F^+ = \{X \to Y \mid F \vdash X \to Y\}$. Dire se una DF appartiene a F^+ è il problema dell'implicazione che tramite l'algoritmo banale (applicando ad F gli assiomi ripetutamente) ha complessità esponenziale.

Chiusura di X rispetto ad F: dato $R<T, F>$ e $X \subseteq T$, la chiusura di X rispetto ad F, denotata con X_F^+ è $X_F^+ = \{A_i \in T \mid F \vdash X \rightarrow A_i\}$

Algoritmo (Chiusura lenta)

```
input R<T, F> e X ⊆ T
output X_F^+
begin
    X_F^+ = X
    while (X_F^+ cambia) do
        for (W → V in F with W ⊆ X_F^+ and V ⋈ X_F^+) do
            X_F^+ = X ∪ V
end
```

Chiave candidata: dato lo schema $R<T, F>$ diremo che un insieme di attributi $W \subseteq T$ è una chiave candidata di R se:

- $W \rightarrow T \in F^+$ (W superchiave)
- $\forall V \subset W,\ V \rightarrow T \notin F^+$ (se $V \subset W$, V non superchiave)

Attributo primo: attributo che appartiene ad almeno una chiave.

Copertura: due insiemi di DF, F e G, sullo schema R sono equivalenti $F \equiv G$, sse $F^+ = G^+$. Se $F \equiv G$, allora F è una copertura di G (e G è copertura di F).

Attributo estraneo: data una $X \rightarrow Y \in F$, si dice che X contiene un attributo estraneo A_i sse $(X - \{A_i\}) \rightarrow Y \in F^+$, cioè $F \vdash (X - \{A_i\}) \rightarrow Y$.

Dipendenza ridondante: $X \to Y$ é una dipendenza ridondante sse $F - \{X \to Y\} \vdash X \to Y$.

Copertura canonica: F é detta copertura canonica sse:

- La parte destra di ogni DF in F é un attributo (1)
- Non esistono attributi estranei (E)
- Nessuna dipendenza in F é ridondante (R)

Decomposizione: dato uno schema $R(T)$, $\rho = \{R(T_1), \dots, R(T_m)\}$ é una decomposizione di R sse $T_1 \cup \dots \cup T_m = T$.

Conservazione dei dati: $\rho = \{R_1(T_1), \dots, R_m(T_m)\}$ é una decomposizione di uno schema $R(T)$ che preserva i dati sse per ogni istanza valida r di R:

$$r = (\pi_{T_1} r) \bowtie \dots \bowtie (\pi_{T_m} r)$$

Uno schema $R(T)$ si decompone senza perdita dei dati negli schemi $R_1(T_1)$ e $R_2(T_2)$ se, per ogni possibile istanza r di $R(T)$, il join naturale delle proiezioni di r su T_1 e T_2 produce la tabelle di partenza.

Proiezione delle dipendenze: dato lo schema $R<T,F>$, e $T_1 \subseteq T$, la proiezione di F su T_1 é $\pi_{T_1}(F) = \{X \to Y \in F^+ \mid XY \subseteq T_1\}$.

Algoritmo (Proiezione)

for each $Y \subseteq T_1$ do
    ($Z = Y^+$; output $Y \to Z \cap T_1$)

Conservazione delle dipendenze: dato lo schema $R<T, F>$, la decomposizione $\rho = \{R_1(T_1), \dots, R_n(T_n)\}$ preserva le dipendenze sse l'unione delle dipendenze in $\pi_{T_i}(F)$ è una copertura di F.

Teorema (Conservazione dei dati): sia $\rho = \{R_i<T_i, F_i>\}$ una decomposizione di $R<T, F>$ che preserva le dipendenze e tale che T_j, per qualche j, è una superchiave per $R<T, F>$. Allora ρ preserva i dati.

BCNF: una relazione r è in forma normale di Boyce Codd se, per ogni dipendenza funzionale $X \to Y$ definita su di essa, X contiene una chiave K di r (è una superchiave).

Algoritmo (Analisi)

input: $R<T, F>$ (F copertura canonica)
output: decomposizione in BCNF che preserva i dati
begin
  $\rho = \{R<T, F>\}$
  while (esiste in ρ una $R_i<T_i, F_i>$ non in BCNF per $X \to A$) do
    $T_a = X^+$
    $F_a = \pi_{T_a}(F_i)$
    $T_b = T_i - X^+ + X$
    $F_b = \pi_{T_b}(F_i)$
    $\rho = \rho - R_i + \{R_a<T_a, F_a>, R_b<T_b, F_b>\}$
end

Teorema (3NF): $R<T, F>$ è in 3NF se per ogni $X \to A \in F^+$, con $A \notin X$, X è una superchiave o A è primo.

Algoritmo (Sintesi)

input: un insieme R di attributi e un insieme F di dipendenze su R
output: una decomposizione $\rho = \{S_i\}$ di R che preservi dati e dipendenze e ogni S_i sia in 3NF, rispetto alle proiezioni di F su S_i.

begin

Passo 1: Trova una copertura canonica G di F e poni $\rho = \{\}$.
Passo 2: Sostituisci in G ogni insieme $X \rightarrow A_1, \ldots, X \rightarrow A_n$ con $X \rightarrow A_1, \ldots, A_n$.
Passo 3: Per ogni $X \rightarrow Y$ in G, metti uno schema con attributi XY in ρ.
Passo 4: Elimina ogni schema di ρ contenuto in un altro schema di ρ.
Passo 5: Se la decomposizione non contiene alcuno schema i cui attributi costituiscano una superchiave per R, aggiungi ad essa lo schema con attributi W, con W chiave di R.

end

Chiave: un insieme di attributi che identifica le n-uple di una relazione.

Chiave primaria: permette di identificare univocamente un'istanza e non può valere NULL.

Chiave esterna: permette di effettuare una relazione tra due tabelle, associando un campo di una tabella al campo chiave primaria di un'altra tabella (vincolo referenziale).

Superchiave: un insieme di attributi che identifica in modo univoco ogni tupla di una relazione e può valere NULL. Una superchiave è anche chiave se è minimale, ovvero non ulteriormente scomponibile.

Superchiave minimale: una superchiave non ulteriormente scomponibile.