Lezione 8 – Chiusura di un insieme di dipendenze funzionali

Prof.ssa Maria De Marsico
demarsico@di.uniroma1.it

SAPIENZA
UNIVERSITÀ DI ROMA

Parte del materiale è tratto dal materiale preparato dalla prof.ssa Moscarini per il canale A-L

Introduciamo F^A

- Ricordiamo che il nostro problema è calcolare l'insieme di dipendenze F^+ che **viene soddisfatto** da **ogni istanza legale** di uno schema *R* su cui è definito un insieme di dipendenze funzionali *F*
- Abbiamo concluso che banalmente $F \subseteq F^+$ in quanto una istanza è **legale solo se** soddisfa **tutte** le dipendenze in *F*
- E le altre?
- Partiamo da un insieme diverso, «facile» da calcolare anche se ... richiede tempo!
- Introduciamo F^A

Assiomi di Armstrong

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

• Denotiamo con F^A l'insieme di dipendenze funzionali definito nel modo seguente:

- se $f \in F$ allora $f \in F^A$
- se $Y \subseteq X \subseteq R$ allora $X \rightarrow Y \in F^A$ (**assioma della riflessività**)
- se $X \rightarrow Y \in F^A$ allora $XZ \rightarrow YZ \in F^A$, per ogni $Z \subseteq R$ (**assioma dell'aumento**)
- se $X \rightarrow Y \in F^A$ e $Y \rightarrow Z \in F^A$ allora $X \rightarrow Z \in F^A$ (**assioma della transitività**)

• **Dimostreremo che $F^+ = F^A$**, cioè che la chiusura di un insieme di dipendenze funzionali F può essere ottenuta a partire da F applicando ricorsivamente gli assiomi della riflessività, dell'aumento e della transitività, conosciuti come **assiomi di Armstrong**.

Qualche semplice osservazione

Sapienza Università di Roma
Dipartimento di Informatica

- se $Y \subseteq X \subseteq R$ allora $X \rightarrow Y \in F^A$ (**assioma della riflessività**)

•Nome $\subseteq$ *(Nome, Cognome)* quindi **ovviamente se** due tuple hanno uguale la coppia *(Nome, Cognome)* ***allora*** avranno sicuramente uguale l'attributo *Nome* (idem per *Cognome), quindi (Nome, Cognome)* $\rightarrow$ *Nome* viene sempre soddisfatta

- se $X \rightarrow Y \in F^A$ allora $XZ \rightarrow YZ \in F^A$, per ogni $Z \subseteq R$ (**assioma dell'aumento**)

CodFiscale $\rightarrow$ *Cognome* è soddisfatta quando, **se** due tuple hanno *CodFiscale* uguale, **allora** hanno anche Cognome uguale.

Se la dipendenza è soddisfatta, e aggiungo l'attributo *Indirizzo,* avrò che **se** due tuple sono uguali su (*CodFiscale, Indirizzo),* **lo devono essere** anche su (*Cognome, Indirizzo) (Indirizzo* è **incluso** nella porzione di tuple che è **uguale**) , quindi se viene soddisfatta

CodFiscale $\rightarrow$ *Cognome*

viene soddisfatta anche

CodFiscale, *Indirizzo* $\rightarrow$ *Cognome, Indirizzo*

Qualche semplice osservazione

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

- se $X \rightarrow Y \in F^A$ e $Y \rightarrow Z \in F^A$ allora $X \rightarrow Z \in F^A$ (**assioma della transitività**)

Matricola→ *CodFiscale* è soddisfatta quando, **se** due tuple hanno *Matricola* uguale, **allora** hanno anche *CodFiscale* uguale.

CodFiscale → *Cognome* è soddisfatta quando, **se** due tuple hanno *CodFiscale* uguale, **allora** hanno anche *Cognome* uguale

Allora se entrambe le dipendenze sono soddisfatte, e due tuple hanno *Matricola* uguale, allora hanno anche *CodFiscale* uguale, ma allora hanno anche *Cognome* uguale, quindi .. . Se entrambe le dipendenze sono soddisfatte, ogni volta che due tuple hanno *Matricola* uguale avranno anche *Cognome* uguale, e quindi viene soddisfatta anche

Matricola→ *Cognome*

Prima di procedere ...

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

- Prima di procedere introduciamo altre tre regole conseguenza degli assiomi che consentono di derivare da dipendenze funzionali in F^A altre dipendenze funzionali in F^A.

- se $X \rightarrow Y \in F^A$ e $X \rightarrow Z \in F^A$ allora $X \rightarrow YZ \in F^A$ (**regola dell'unione**)

- se $X \rightarrow Y \in F^A$ e $Z \subseteq Y$ allora $X \rightarrow Z \in F^A$ (**regola della decomposizione**)

- se $X \rightarrow Y \in F^A$ e $WY \rightarrow Z \in F^A$ allora $WX \rightarrow Z \in F^A$ (**regola della pseudotransitività**).

Dimostriamolo

SAPIENZA
Università di Roma
Dipartimento di Informatica

•**Teorema** Sia F un insieme di dipendenze funzionali. Valgono le seguenti implicazioni:

a) se $X \rightarrow Y \in F^A$ e $X \rightarrow Z \in F^A$ allora $X \rightarrow YZ \in F^A$

b) se $X \rightarrow Y \in F^A$ e $Z \subseteq Y$ allora $X \rightarrow Z \in F^A$

c) se $X \rightarrow Y \in F^A$ e $WY \rightarrow Z \in F^A$ allora $WX \rightarrow Z \in F^A$.

•**Dim.**

stiamo trattando con **insiemi!** Quindi … XX=X

•**Prova di a).**

•Se $X \rightarrow Y \in F^A$**,** per l'assioma **dell'aumento** si ha $X \rightarrow XY \in F^A$**.** Analogamente**,** se $X \rightarrow Z \in F^A$**,** per l'assioma **dell'aumento** si ha $XY \rightarrow YZ \in F^A$**.** Quindi, poiché $X \rightarrow XY \in F^A$ e $XY \rightarrow YZ \in F^A$, per l'assioma della **transitività** si ha $X \rightarrow YZ \in F^A$.

•**Prova di b).**

•Se $Z \subseteq Y$ allora, per l'assioma della **riflessività**, si ha $Y \rightarrow Z \in F^A$. Quindi, poiché $X \rightarrow Y \in F^A$ e $Y \rightarrow Z \in F^A$. per l'assioma della **transitività** si ha $X \rightarrow Z \in F^A$.

•**Prova di c).**

•Se $X \rightarrow Y \in F^A$, per l'assioma **dell'aumento** si ha $WX \rightarrow WY \in F^A$**.** Quindi, poiché $WX \rightarrow WY \in F^A$ e $WY \rightarrow Z \in F^A$, per l'assioma della **transitività** si ha $WX \rightarrow Z \in F^A$.

□

Osservazione

Osserviamo che:

- per la regola dell'**unione**, se $X \rightarrow A_i \in F^A$, $i=1, \ldots, n$ allora $X \rightarrow A_1, \ldots, A_i \ldots A_n \in F^A$

- per la regola della **decomposizione**, *se* $X \rightarrow A_1, \ldots, A_i \ldots A_n \in F^A$ *allora* $X \rightarrow A_i \in F^A$, $i=1, \ldots, n$

quindi

se e solo se

- $X \rightarrow A_1, \ldots, A_i \ldots A_n \in F^A \Leftrightarrow X \rightarrow A_i \in F^A$, $i=1, \ldots, n$

e possiamo limitarci in generale a considerare le dipendenze col membro destro singleton

Chiusura di un insieme di attributi

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

- ***Definizione***

Siano *R* uno schema di relazione, *F* un insieme di dipendenze funzionali su *R* e *X* un sottoinsieme di *R*.

La *chiusura* di *X* rispetto ad *F*, denotata con X^+_F (o semplicemente X^+, se non sorgono ambiguità) è definito nel modo seguente:

$$X^+_F = \{A \mid X \rightarrow A \in F^A\}$$

- In pratica fanno parte della chiusura di un insieme di attributi X tutti quelli che **sono determinati** funzionalmente da X eventualmente **applicando gli assiomi di Armstrong**

Banalmente … $\boldsymbol{X \subseteq X^+_F}$ (riflessività!)

Lemma

Sapienza Università di Roma
Dipartimento di Informatica

- **Lemma** Siano R uno schema di relazione ed F un insieme di dipendenze funzionali su R. Si ha che: $X \to Y \in F^A$ se e solo se $Y \subseteq X^+$.
- **Dim.**

Sia $Y = A_1, A_2, \ldots, A_n$.

- **<u>Parte se</u>.**
- Poiché $Y \subseteq X^+$, per ogni i, $i = 1, \ldots, n$, si ha che $X \to A_i \in F^A$. Pertanto, per la regola **dell'unione**, $X \to Y \in F^A$.
- **<u>Parte solo se</u>.**
- Poiché $X \to Y \in F^A$, per la regola della **decomposizione** si ha che, per ogni i, $i = 1, \ldots, n$, $X \to A_i \in F^A$, cioè $A_i \in X^+$, per ogni i, i$= 1, \ldots, n$, e, quindi, $Y \subseteq X^+$.

Teorema: $F^+=F^A$

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

- **Teorema** Siano *R* uno schema di relazione ed *F* un insieme di dipendenze funzionali su *R*. Si ha $F^+ = F^A$.
- **Dim.** (l'uguaglianza d insiemi di dimostra dimostrando la **doppia inclusione**)

contiene

- $\boldsymbol{F^+ \supseteq F^A}$**.** Sia $X \rightarrow Y$ una dipendenza funzionale in F^A. Dimostriamo che $X \rightarrow Y \in F^+$ **per induzione** sul numero ***i*** di **applicazioni di uno degli assiomi di Armstrong**.

Base dell'induzione: $i=0$. In tal caso $X \rightarrow Y$ è in *F* e quindi, banalmente, $X \rightarrow Y$ è in F^+.

Induzione: $i>0$. Per **l'ipotesi induttiva** ogni dipendenza funzionale ottenuta a partire da *F* applicando gli assiomi di Armstrong un numero di volte **minore o uguale a *i–1*** è in F^+. Dobbiamo dimostrarlo per un numero di volte uguale a ***i***. Si possono presentare tre casi

Teorema: $F^+=F^A$ (segue)

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

- $X \rightarrow Y$ è stata ottenuta mediante l'assioma della **riflessività**, in tal caso $Y \subseteq X$. Sia r un'istanza di R e siano t_1 e t_2 due tuple di r tali che $t_1[X]= t_2[X]$; banalmente si ha $t_1[Y]=t_2[Y]$.

Teorema: $F^+=F^A$ (segue)

- $X \rightarrow Y$ è stata ottenuta applicando l'assioma **dell'aumento** ad una dipendenza funzionale $V \rightarrow W$ **in** F^A, ottenuta a sua volta applicando ricorsivamente gli assiomi di Armstrong un numero di volte **minore o uguale a** $i-1$ (quindi **per l'ipotesi induttiva** $V \rightarrow W \in F^+$); sarà quindi $X=VZ$ **e** $Y=WZ$, **per qualche** $Z \subseteq R$. Sia r un'istanza **legale** di R e siano t_1 e t_2 due tuple di r tali che $t_1[X]=t_2[X]$; banalmente si ha **che** $t_1[V]=t_2[V]$ e $t_1[Z]=t_2[Z]$. Per **l'ipotesi induttiva** da $t_1[V]=t_2[V]$ **segue** $t_1[W]=t_2[W]$; da $t_1[W]=t_2[W]$ **e** $t_1[Z]=t_2[Z]$ **segue** $t_1[Y]=t_2[Y]$.

Teorema: $F^{+}=F^{A}$ (segue)

- $X \rightarrow Y$ è stata ottenuta applicando l'assioma della transitività a due dipendenze funzionale $X \rightarrow Z$ e $Z \rightarrow Y$ in F^{A}, ottenute a loro volta applicando ricorsivamente gli assiomi di Armstrong un numero di volte minore o uguale a $i-1$ (quindi per l'ipotesi induttiva **$X \rightarrow Z$ e $Z \rightarrow Y \in F^{+}$**). Sia r un'istanza **legale** di R e siano t_1 e t_2 due tuple di r tali che $t_1[X]=t_2[X]$. Per l'ipotesi induttiva da **$t_1[X]=t_2[X]$** segue $t_1[Z]=t_2[Z]$; da $t_1[Z]=t_2[Z]$, ancora per l'ipotesi induttiva segue **$t1[Y]=t_2[Y]$**.

Teorema: $F^+=F^A$ (segue)

•$\boldsymbol{F^+ \subseteq F^A}$**.** Supponiamo **per assurdo** che esista una dipendenza funzionale $X \to Y \in F^+$ tale che $\boldsymbol{X \to Y \notin F^A}$. Useremo una particolare istanza **legale** di *R* per dimostrare che questa supposizione porta ad una contraddizione.

Consideriamo la seguente istanza *r* di *R*:

	X^+				$R\text{-}X^+$			
r	**1**	**1**	**…**	**1**	**1**	**1**	**…**	**1**
	1	**1**	**…**	**1**	**0**	**0**	**…**	**0**

L’istanza da solo **due** tuple, **uguali** sugli attributi in $\boldsymbol{X^+}$ e **diverse** in tutti gli altri ($\boldsymbol{R\text{-}X^+}$)

Dimostreremo che l’istanza è **legale**
Utilizzando questa istanza dimostreremo che se $X \to Y \in F^+$ **NON** può succedere che $\boldsymbol{X \to Y \notin F^A}$ (assumendo il contrario arriviamo ad una contraddizione).

Teorema: $F^+ = F^A$ (segue)

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

- Mostriamo che:
  - ***r*** **è un'istanza legale di** ***R***. Sia $V \rightarrow W$ una dipendenza funzionale **in** ***F*** e supponiamo **per assurdo** che **non sia soddisfatta** da *r* (*r* non è legale). In tal caso le **due tuple** di *r* devono avere **gli stessi valori per** ***V*** **e differenti valori per** ***W***; ciò implica che $V \subseteq X^+$ e $W \cap (R-X^+) \neq \emptyset$. Poiché $V \subseteq X^+$, **per il Lemma, si ha che** $X \rightarrow V \in F^A$; pertanto, **per l'assioma della transitività** $(X \rightarrow V$ **e** $V \rightarrow W)$, $X \rightarrow W \in F^A$ e, quindi, **di nuovo per il Lemma ,** $W \subseteq X^+$ (che contraddice $W \cap (R- X^+) \neq \emptyset$). **Quindi** $V \rightarrow W$ **è soddisfatta.** Quindi siamo arrivati **ad una contraddizione**. Poiché abbiamo considerato una **qualunque** dipendenza in *F*, *r* le soddisfa tutte, quindi è **legale.** □

per avere **gli stessi valori** sulle due tuple **TUTTI** gli attributi di V devono essere in X^+
per avere **valori diversi** sulle due tuple **ALMENO UN** attributo di W deve stare fuori da X^+

Teorema: $F^+=F^A$ (segue)

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

- Mostriamo che:

  $X \rightarrow Y \in F^+ \Rightarrow X \rightarrow Y \in F^A$. Supponiamo per assurdo che $X \rightarrow Y \in F^+$ e $X \rightarrow Y \notin F^A$. Abbiamo mostrato che la nostra *r* è un'istanza legale. Poiché le dipendenze in F^+ sono soddisfatte da ogni istanza legale, allora *r* soddisfa $X \rightarrow Y$. ***Abbiamo due tuple uguali su X? Certo!*** Poiché $X \subseteq X^+$ (per l'assioma della riflessività e il Lemma) le due tuple di *r* coincidono sugli attributi *X* e quindi, poiché *r* soddisfa $X \rightarrow Y$, devono coincidere anche sugli attributi di *Y*. Questo implica che $Y \subseteq X^+$ e, per il Lemma, che $X \rightarrow Y \in F^A$, arrivando ad una <u>contraddizione</u>. □

Teorema: $F^+ = F^A$ (Nota finale)

- E' utile notare che la dimostrazione di questo teorema si basa su due collegamenti MOLTO importanti:
  - Il collegamento che esiste tra l'insieme di dipendenze F^+ e le istanze legali: **se** una istanza è legale **allora** soddisfa anche tutte le dipendenze in F^+ e d'altra parte F^+ è l'insieme di dipendenze soddisfatte **da ogni** istanza legale (notare che per verificare se una istanza è legale basta controllare che soddisfi le dipendenze in F)

  - Il collegamento che esiste **tra la chiusura X^+** di un insieme di attributi X (diamo per scontato che l'insieme di dipendenze di riferimento sia F e omettiamo il pedice) e il **sottoinsieme di dipendenze in F^A** (che abbiamo appena visto essere uguale a F^+) **che hanno X come determinante**, cioè $Y \subseteq X^+ \Leftrightarrow X \rightarrow Y \in F^A$ che equivale a dire che $Y \subseteq X^+ \Leftrightarrow X \rightarrow Y \in F^+$ e in particolare che $X \rightarrow Y$ **deve essere soddisfatta da ogni istanza legale**□

A cosa ci serve conoscere F^+?

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

- Quindi abbiamo ora un modo per identificare tutte le dipendenze in F^+. Sono esattamente le stesse che si possono inserire in F^A partendo da F e applicando gli assiomi di Armstrong e le regole derivate (**a costo però di una complessità esponenziale**)

- Calcolare $\boldsymbol{F^A}$, e quindi F^+, richiede tempo esponenziale in $|R|$. Basta considerare gli assiomi della riflessività e dell'aumento oppure la regola della decomposizione; ogni possibile sottoinsieme di R porta a una o più dipendenze … e i possibili sottoinsiemi di R sono $\boldsymbol{2^{|R|}}$ … pertanto il calcolo di $|F^+|$ ha complessità esponenziale in $|R|$.

A cosa ci serve conoscere F^+?

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

- Parleremo di *terza forma normale* (*3NF),* delle sue proprietà e di come ottenerla se lo schema di relazione che abbiamo progettato è carente. Come abbiamo già accennato, il problema nasce dal fatto di aver rappresentato più concetti (oggetti) nella stessa relazione, e che la soluzione consiste nel **decomporre** lo schema in maniera opportuna.

- Ovviamente, **tutte** le dipendenze funzionali **originarie** devono essere soddisfatte **anche** dalle istanze dei nuovi schemi. **In particolare, devono essere preservate tutte le dipendenze in F^+.**