

BASI DI DATI

DIPENDENZE FUNZIONALI E FORME NORMALI

Prof. Fabio A. Schreiber

Dipartimento di Elettronica e Informazione
Politecnico di Milano



ERRORI DI PROGETTAZIONE

- **INSERIMENTO DI ELEMENTI RIDONDANTI**
 - SPRECO DI MEMORIA
 - VIOLAZIONE DELL'INTEGRITA' DI UNA RELAZIONE A SEGUITO DI AGGIORNAMENTO DI UNA TUPLA
(ANOMALIA DA AGGIORNAMENTO)
- **IMPOSSIBILITA' DI RAPPRESENTARE ALCUNE INFORMAZIONI**
 - UTILIZZO **SELVAGGIO** DI VALORI NULLI **(ANOMALIE DA INSERIMENTO/CANCELLAZIONE)**
- **PERDITA DI INFORMAZIONI**
 - IMPOSSIBILITA' DI ESEGUIRE CORRETTAMENTE I JOIN
(NASCITA DI TUPLE SPURIE)

FORME NORMALI

PERMETTONO AL PROGETTISTA DI

- ANALIZZARE GLI SCHEMI DI RELAZIONE BASANDOSI SULLE LORO CHIAVI E SULLE DIPENDENZE FUNZIONALI TRA GLI ATTRIBUTI
- ESEGUIRE DEI TEST PER VERIFICARE L'APPARTENENZA DI UNA RELAZIONE AD UNA FORMA NORMALE
 - SE IL TEST FALLISCE, LA RELAZIONE DEVE ESSERE DECOMPOSTA IN PIU' RELAZIONI CHE INDIVIDUALMENTE LO SODDISFANO

© Fabio A. Schreiber

Progettazione 2

FORME NORMALI

- APPLICANDOSI A SINGOLE RELAZIONI, LE FORME NORMALI **DA SOLE NON GARANTISCONO** DI EVITARE TUTTE LE ANOMALIE
- E' NECESSARIO QUINDI VERIFICARE ANCHE
 - LA PROPRIETA' DI POTER ESEGUIRE **JOIN SENZA PERDITA**
 - GLI ATTRIBUTI SUI QUALI VIENE ESEGUITO IL JOIN DEVONO ESSERE CHIAVI O CHIAVI ESTERNE
 - LA PROPRIETA' DI **CONSERVAZIONE DELLE DIPENDENZE FUNZIONALI** (EVITA LE ANOMALIE DI INSERIMENTO/ CANCELLAZIONE)
 - OGNI DIPENDENZA FUNZIONALE DEVE ESSERE PRESENTE IN ALMENO UNA RELAZIONE

© Fabio A. Schreiber

Progettazione 3

ELEMENTI RIDONDANTI

- ESEMPIO 1

MUTUO

(numero, **contraente**, **valore**, filiale, **patrimonio**, **città**)

- PATRIMONIO E CITTA' SONO ATTRIBUTI COLLEGATI A FILIALE
- CONTRAENTE E VALORE SONO ATTRIBUTI COLLEGATI A NUMERO
- PER OGNI MUTUO STIPULATO IN UNA FILIALE BISOGNA RIPETERE I VALORI DI PATRIMONIO E CITTA'
- OGNI VOLTA CHE CAMBIA IL VALORE DEL PATRIMONIO (ABBASTANZA SPESSO) BISOGNA CAMBIARLO IN TUTTE LE TUPLE DELLA STESSA FILIALE

- ESEMPIO 2

BARCA

(matr, nome, anno, modello, **lunghezza**, **stazza**, armatore)

- LUNGHEZZA E STAZZA SONO ATTRIBUTI COLLEGATI A MODELLO
- PER OGNI BARCA DELLO STESSO MODELLO BISOGNA RIPETERE I VALORI DI LUNGHEZZA E STAZZA

© Fabio A. Schreiber

Progettazione 4

RAPPRESENTAZIONE DI INFORMAZIONI

- ESEMPIO

RAPPORTO

(correntista, filiale, num_conto, saldo, num_mutuo, valore)

- IN QUESTO SCHEMA NON E' POSSIBILE RAPPRESENTARE CORRENTISTI CHE NON ABBIANO CONTRATTO MUTUI SE NON INTRODUCENDO **VALORI NULLI** PER num_mutuo E valore

© Fabio A. Schreiber

Progettazione 5

PERDITA DI INFORMAZIONI

- **ESEMPIO**

VALORE_PRESTITO (correntista, valore)

MUTUO (numero, filiale, valore)

- NON E' POSSIBILE RISALIRE CORRETTAMENTE AL NUMERO DEL MUTUO E ALLA FILIALE PRESSO LA QUALE IL CORRENTISTA LO DETIENE
 - IL valore DIPENDE DAL CORRENTISTA E DIPENDE DAL NUMERO DEL MUTUO, MA CI POSSONO ESSERE MOLTI CORRENTISTI CHE HANNO MUTUI CON LO STESSO VALORE
 - NEL FARE IL JOIN VENGONO GENERATE TUTTE LE POSSIBILI TUPLE CHE SONO LA COMBINAZIONE DI MUTUI E CORRENTISTI CHE PRESENTANO LO STESSO valore
 - MOLTE DI QUESTE TUPLE NON HANNO UNA CORRISPONDENZA CON LA REALTA', MA NON E' POSSIBILE INDIVIDUARLE → IL JOIN E' **LOSSY**

© Fabio A. Schreiber

Progettazione 6

DECOMPOSIZIONE DI RELAZIONI

Siano U, R_1, \dots, R_n schemi di relazioni tali che

$$\bigcup_{i=1}^n R_i = U$$

Gli schemi R_i costituiscono una **Decomposizione** di U

© Fabio A. Schreiber

Progettazione 7

DECOMPOSIZIONE DI RELAZIONI

Una relazione è **legale** se soddisfa tutte le regole e i vincoli **C** che vengono imposti sulla Base di Dati

Una decomposizione $\{R_1, \dots, R_n\}$ di uno schema U è **lossless-join** se, per tutte le relazioni u su U legali

$$u = \bowtie_1^n \prod_{R_i} (u)$$

© Fabio A. Schreiber

Progettazione 8

DIPENDENZE FUNZIONALI

- SONO UNA **GENERALIZZAZIONE** DEL CONCETTO DI **CHIAVE**
- PERMETTONO DI **ESPRIMERE** UN INSIEME DI **VINCOLI** TRA GLI ATTRIBUTI IN UNA BASE DI DATI
- PORTANO ALLA **DEFINIZIONE** DI ALCUNE **FORME NORMALI** CHE GARANTISCONO LA BONTA' DEL PROGETTO

© Fabio A. Schreiber

Progettazione 9

DIPENDENZE FUNZIONALI

SIA R UNO SCHEMA DI RELAZIONE E SIANO V E W DUE SOTTOINSIEMI (STRINGHE) DI ATTRIBUTI DI R

$$V \subseteq R; \quad W \subseteq R$$

SI DICE CHE $V \rightarrow W$ (V DETERMINA W O W DIPENDE FUNZIONALMENTE DA V) SE, **PER OGNI COPPIA** DI TUPLE t_1, t_2 APPARTENENTI ALLA RELAZIONE CON SCHEMA R ,

$$t_1[V]=t_2[V] \Rightarrow t_1[W]=t_2[W]$$

NEL CASO DI UNA SUPERCHIAVE $K \Rightarrow R$

$$t_1[K]=t_2[K] \Rightarrow t_1=t_2$$

DIPENDENZE BANALI

$V \rightarrow W$ E' BANALE SSE $W \subseteq V$

CASO PARTICOLARE $A \rightarrow A$

© Fabio A. Schreiber

Progettazione 10

DIPENDENZE FUNZIONALI

USO DELLE DF

- SPECIFICARE VINCOLI SULL'INSIEME DI RELAZIONI LEGALI
- VERIFICARE SE UN INSIEME DI RELAZIONI E' LEGALE **RISPETTO AD UN INSIEME** DI DF



- DETERMINARE **TUTTE** LE DF
- SCEGLIERE UNA DECOMPOSIZIONE
- VERIFICARE SE L'INSIEME DI RELAZIONI RISULTANTI E' LEGALE RISPETTO ALL'INSIEME DI DF

© Fabio A. Schreiber

Progettazione 11

DIPENDENZE FUNZIONALI

DA DOVE DERIVANO LE *DF*?

- LE *DF* DOVREBBERO DERIVARE DIRETTAMENTE DALL'ANALISI DEI REQUISITI
- E' POSSIBILE **INFERIRE** LE *DF* VALIDE SU UN INSIEME DI ATTRIBUTI ESAMINANDO UN INSIEME DI **ISTANZE DI RELAZIONI**
 - **ATTENZIONE** E' INDISPENSABILE **VALIDARE** CON LE SPECIFICHE FUNZIONALI LE *DF* OTTENUTE IN QUESTO MODO E SOPRATTUTTO **USARE IL BUON SENSO**

© Fabio A. Schreiber

Progettazione 12

ESEMPIO DI INFERENZA DI *DF*

A	B	C	D
a ₁	b ₁	c ₁	d ₁
a ₁	b ₂	c ₁	d ₂
a ₂	b ₂	c ₂	d ₂
a ₂	b ₃	c ₂	d ₃
a ₃	b ₃	c ₂	d ₄

$a, b_j = \alpha_k$

α	C	D
α_1	c ₁	d ₁
α_2	c ₁	d ₂
α_3	c ₂	d ₂
α_4	c ₂	d ₃
α_5	c ₂	d ₄

$A \rightarrow C$

C non determina A

$AB \rightarrow D$

© Fabio A. Schreiber

Progettazione 13

DIPENDENZE FUNZIONALI

SUGLI ESEMPI PRECEDENTI

- IN MUTUO
numero, nome_filiale → nome_contraente, valore
nome_filiale → patrimonio, città
- IN BARCA
matr → nome, anno, armatore, modello
modello → lunghezza, stazza
- IN RAPPORTO
num_conto → correntista, saldo
num_mutuo → correntista, valore
- IN VALORE_PRESTITO
correntista → valore

© Fabio A. Schreiber

Progettazione 14

CHIUSURA DELLE *DF*

GLI ALGORITMI PER LA
DECOMPOSIZIONE DI RELAZIONI
RICHIEDONO LA CONOSCENZA DI
TUTTE LE *DF* DEFINITE PER IL SISTEMA

© Fabio A. Schreiber

Progettazione 15

CHIUSURA DELLE DF

ASSIOMI DI ARMSTRONG

INSIEME **CORRETTO E COMPLETO** CHE GENERA TUTTE LE DF IN \mathcal{F}^+

- RIFLESSIVITÀ: $Y \subseteq X \Rightarrow X \rightarrow Y$
- AUMENTO: $X \rightarrow Y \Rightarrow WX \rightarrow WY$
- TRANSITIVITÀ: $X \rightarrow Y \text{ E } Y \rightarrow Z \Rightarrow X \rightarrow Z$

ALTRE REGOLE UTILI

- UNIONE: $X \rightarrow Y \text{ E } X \rightarrow Z \Rightarrow X \rightarrow YZ$
- DECOMPOSIZIONE: $X \rightarrow YZ \Rightarrow X \rightarrow Y \text{ E } X \rightarrow Z$
- PSEUDOTRANSITIVITÀ: $X \rightarrow Y \text{ E } WY \rightarrow Z \Rightarrow WX \rightarrow Z$

© Fabio A. Schreiber

Progettazione 16

CHIUSURA DELLE DF

L'INSIEME DI TUTTE LE DF DERIVABILI DA UN INSIEME \mathcal{F} DI DF DI PARTENZA MEDIANTE GLI ASSIOMI DI ARMSTRONG E' DETTO **CHIUSURA** E SI INDICA CON \mathcal{F}^+

NEL CASO PEGGIORE, L'ALGORITMO CHE CALCOLA \mathcal{F}^+ E' ESPONENZIALE E COMUNQUE LA COMPLESSITA' PER UN CALCOLO DIRETTO A PARTIRE DALLA DEFINIZIONE DI DF E' TROPPO ELEVATA PER UN UTILIZZO PRATICO

© Fabio A. Schreiber

Progettazione 17

ESEMPI

SIANO DATI:

- UNO SCHEMA $R(A,B,C,G,H,I)$
 - UN INSIEME DI DF $\mathcal{F}=\{A\rightarrow B, A\rightarrow C, CG\rightarrow H, CG\rightarrow I, B\rightarrow H\}$
- RISULTANO ANCHE: $A\rightarrow H, CG\rightarrow HI, AG\rightarrow I, \dots$

ESEMPIO DI PROVA DIRETTA

$$A\rightarrow B \Leftrightarrow (t_1[A]=t_2[A] \Rightarrow t_1[B]=t_2[B])$$

$$B\rightarrow H \Leftrightarrow (t_1[B]=t_2[B] \Rightarrow t_1[H]=t_2[H])$$

$$\Rightarrow A\rightarrow H$$

ESEMPI USANDO LE REGOLE

- $A\rightarrow H$ TRANSITIVITÀ $A\rightarrow B, B\rightarrow H$
- $CG\rightarrow HI$ UNIONE $CG\rightarrow H, CG\rightarrow I$
- $AG\rightarrow I$ AUMENTO $A\rightarrow C \Rightarrow AG\rightarrow CG$ +
TRANSITIVITÀ $CG\rightarrow I$

© Fabio A. Schreiber

Progettazione 18

ESEMPI

SPESSO BASTA DETERMINARE SOLAMENTE QUALI ATTRIBUTI SIANO DETERMINATI DA UN INSIEME DI ATTRIBUTI X IN BASE ALLE DIPENDENZE DI \mathcal{F} (CHIUSURA X^+ DI X SOTTO \mathcal{F})

- L'INSIEME DI PARTENZA FA PARTE DEL RISULTATO
- ALL'INSIEME DI PARTENZA SI AGGIUNGE MAN MANO LA PARTE DESTRA DELLE DF LA CUI PARTE SINISTRA COMPARE GIÀ NEL RISULTATO
- IL PROCEDIMENTO TERMINA QUANDO NON CI SONO PIU' DF CHE SODDISFANO LA PROPRIETÀ

© Fabio A. Schreiber

Progettazione 19

ESEMPI

NELL'ESEMPIO PRECEDENTE CALCOLIAMO $(AG)^+$

AG

$AG \rightarrow B$ $AG \cup B$

AGB

$A \rightarrow C$ $AGB \cup C$

AGBC

$CG \rightarrow H$ $AGBC \cup H$

AGBCH

$CG \rightarrow I$ $AGBCH \cup I$

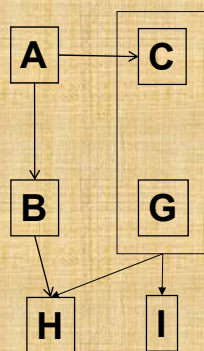
AGBCHI

QUINDI AG DETERMINANO TUTTI GLI ALTRI ATTRIBUTI
DELLO SCHEMA E PERTANTO NE SONO UNA **CHIAVE**

© Fabio A. Schreiber

Progettazione 20

RAPPRESENTAZIONE GRAFICA DI DF



$R_1(\underline{A}, B, C)$
 $R_2(\underline{C}, G, H, I)$
 $R_3(\underline{B}, H)$

LE PROPRIETA' DI INTERCONNESSIONE DEL GRAFO FORNISCONO
UTILI INDICAZIONI PER LA NORMALIZZAZIONE DELLE RELAZIONI

© Fabio A. Schreiber

Progettazione 21

USO DELLE DF

1. VERIFICA DI DECOMPOSIZIONE LOSSLESS-JOIN

CRITERIO

- SIANO R_1, R_2 UNA DECOMPOSIZIONE DI R
- PERCHE' ESSA SIA LOSSLESS-JOIN **ALMENO UNA** DELLE SEGUENTI DF DEVE ESSERE IN \mathcal{F}^+

- $R_1 \cap R_2 \rightarrow R_1$
- $R_1 \cap R_2 \rightarrow R_2$

© Fabio A. Schreiber

Progettazione 22

ESEMPIO 1

CORRENTISTA (nome, via, città)

FILIALE (nome, patrimonio, città)

PRESTITO (nome_filiale, numero, nome_correntista, valore)

DEPOSITO (nome_filiale, numero, nome_correntista, saldo)

MUTUO (nome_filiale, patrimonio, città, numero, nome_contraente, valore)

$\mathcal{F}_{\text{CORRENTISTA}} \{ \text{nome} \rightarrow \text{città}; \text{nome} \rightarrow \text{via} \}$

$\mathcal{F}_{\text{FILIALE}} \{ \text{nome} \rightarrow \text{città}; \text{nome} \rightarrow \text{patrimonio} \}$

$\mathcal{F}_{\text{PRESTITO}} \{ \text{numero_prestito} \rightarrow \text{valore}; \text{numero_prestito} \rightarrow \text{nome_filiale} \}$

$\mathcal{F}_{\text{DEPOSITO}} \{ \text{numero_deposito} \rightarrow \text{saldo}; \text{numero_deposito} \rightarrow \text{nome_filiale} \}$

$\mathcal{F}_{\text{MUTUO}} \{ \text{numero, nome_filiale} \rightarrow \text{nome_contraente, valore}; \text{nome_filiale} \rightarrow \text{patrimonio, città} \}$

© Fabio A. Schreiber

Progettazione 23

ESEMPIO 1

$R_1(\text{nome_filiale}, \text{patrimonio})$

$R_1'(\text{nome_filiale}, \text{città}, \text{numero}, \text{nome_contraente}, \text{valore})$

$R_1 \cap R_1' = (\text{nome_filiale})$

$\text{nome_filiale} \rightarrow \text{patrimonio}$

$\text{nome_filiale} \rightarrow \text{nome_filiale}, \text{patrimonio}$ per aumento

$\text{nome_filiale}, \text{patrimonio} \equiv R_1$

$R_2(\text{nome_filiale}, \text{città})$

$R_2'(\text{nome_filiale}, \text{numero}, \text{nome_contraente}, \text{valore})$

$R_3(\text{numero}, \text{valore})$

$R_4(\text{numero}, \text{nome_filiale})$

$R_5(\text{numero}, \text{nome_contraente})$

© Fabio A. Schreiber

Progettazione 24

CONSERVAZIONE DELLE DF

- **EVITA LA CREAZIONE DI RELAZIONI ILLEGALI DURANTE LE OPERAZIONI DI AGGIORNAMENTO**
- **IL RISPETTO DELLE DF DEVE POTER ESSERE VERIFICATO GUARDANDO UNA SINGOLA RELAZIONE, SENZA DOVER RICORRERE A JOIN**

© Fabio A. Schreiber

Progettazione 25

USO DELLE DF

2.VERIFICA DI CONSERVAZIONE DELLE DF

SI DEFINISCE **RESTRIZIONE \mathcal{F}_i DI \mathcal{F} A R_i** L'INSIEME DI TUTTE LE DF IN \mathcal{F}^+ CHE CONTENGONO SOLO ATTRIBUTI DI R_i

$$\mathcal{F}' = \bigcup_{i=1}^n \mathcal{F}_i$$

IN GENERALE $\mathcal{F}' \neq \mathcal{F}$

CRITERIO

UNA DECOMPOSIZIONE CONSERVA LE DF SE

$$\mathcal{F}'^+ = \mathcal{F}^+$$

© Fabio A. Schreiber

Progettazione 26

ESEMPIO 2

NELLA DECOMPOSIZIONE DELL'ESEMPIO PRECEDENTE

$R_1(\text{nome_filiale}, \text{patrimonio})$

$R_2(\text{nome_filiale}, \text{città})$

$R_3(\text{numero}, \text{valore})$

$R_4(\text{numero}, \text{nome_filiale})$

$R_5(\text{numero}, \text{nome_contraente})$

$\text{nome_filiale} \rightarrow \text{patrimonio}$

$\text{nome_filiale} \rightarrow \text{città}$

$\text{numero} \rightarrow \text{valore}$

$\text{numero} \rightarrow \text{nome_filiale}$

VIENE VERIFICATA SU R_1

VIENE VERIFICATA SU R_2

VIENE VERIFICATA SU R_3

VIENE VERIFICATA SU R_4

© Fabio A. Schreiber

Progettazione 27

USO DELLE DF

3. ELIMINAZIONE DELLA RIDONDANZA

LA DECOMPOSIZIONE $\{R_1, \dots, R_5\}$

- E' LOSSLESS-JOIN
- CONSERVA LE DF
- NON COMPORTA RIDONDANZA

LA POSSIBILITA' DI OTTENERE LE TRE PROPRIETA' CONTEMPORANEAMENTE **NON** E' SEMPRE GARANTITA, MA DIPENDE DALLA **FORMA NORMALE** NELLA QUALE SI RIESCE A DECOMPORRE LO SCHEMA INIZIALE

© Fabio A. Schreiber

Progettazione 28

FORMA NORMALE DI BOYCE-CODD BCNF

UNO SCHEMA R E' IN **BCNF** SE, **PER TUTTE LE DF** $X \rightarrow Y$ ($X \subseteq R, Y \subseteq R$) CHE VALGONO SU R , VALE **ALMENO UNA** DELLE SEGUENTI CONDIZIONI

- $X \rightarrow Y$ E' BANALE ($Y \subseteq X$)
- X E' UNA SUPERCHIAVE DI R

SI DICE CHE IL PROGETTO DI UNA BASE DI DATI E' IN **BCNF** SE **OGNI SUO SCHEMA DI RELAZIONE E' IN BCNF**

© Fabio A. Schreiber

Progettazione 29

ESEMPIO 3: BCNF

CORRENTISTA (nome, via, città)

E' IN BCNF POICHE' nome E' UNA CHIAVE CANDIDATA E LE *DF* GIA' DEFINITE CONTENGONO nome A SINISTRA DELLA FRECCIA

PRESTITO (nome_filiale, numero, nome_correntista, valore)

NON E' IN BCNF POICHE' numero, PUR ESSENDO A SINISTRA DELLA FRECCIA NELLE *DF*, NON E' CHIAVE PERCHE' NON C'E' UNA *DF* numero → correntista.

INFATTI DUE CORRENTISTI POSSONO AVERE LO STESSO NUMERO DI PRESTITO (O DI CONTO) E IN QUESTA RELAZIONE SI DEVONO RIPETERE LE STESSO INFORMAZIONI PER OGNI TITOLARE DELLO STESSO PRESTITO

© Fabio A. Schreiber

Progettazione 30

ESEMPIO 3: BCNF

POSSIAMO DECOMPORRE LA RELAZIONE:

PRESTITO (nome_filiale, numero, nome_correntista, valore)

IN

- PRESTITO_INFO (numero, nome_filiale, valore)
- PRESTITO_TITOLARE (numero, correntista)

SULLA PRIMA VALGONO LE STESSO *DF* DI PRESTITO

SULLA SECONDA VALGONO SOLO *DF* BANALI

SI PUO' VERIFICARE CHE QUESTA DECOMPOSIZIONE E' ANCHE LOSSLESS-JOIN

© Fabio A. Schreiber

Progettazione 31

DECOMPOSIZIONE IN BCNF

DECOMPOSIZIONE IN BCNF DI UNO
SCHEMA R_i CON $DF X \rightarrow Y$

- $R_1(XY)$
- $R_2(R_i - Y)$

E' LOSSLESS-JOIN, INFATTI

- $(XY) \cap (R_i - Y) = X$; MA $X \rightarrow Y$; $X \rightarrow XY$
 $= R_1$

© Fabio A. Schreiber

Progettazione 32

DECOMPOSIZIONE IN BCNF

ABBIAMO GIA' VISTO CHE LA DECOMPOSIZIONE DELLA
RELAZIONE

MUTUO

(nome_filiale, patrimonio, città, numero, nome_contraente, valore)

IN

R_1 (nome_filiale, patrimonio)

R_2 (nome_filiale, città)

R_3 (numero, valore)

R_4 (numero, nome_filiale)

R_5 (numero, nome_contraente)

- E' LOSSLESS-JOIN
- PRESERVA LE DF

MA CIO' NON ACCADE SEMPRE



© Fabio A. Schreiber

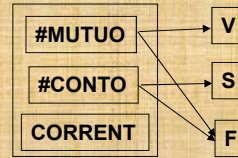
Progettazione 33

ESEMPIO 4

RAPPORTO

(correntista, nome_filiale, num conto, saldo, num mutuo, valore)

DF_1 num_mutuo \rightarrow valore
 DF_2 num_mutuo \rightarrow nome_filiale
 DF_3 num_conto \rightarrow saldo
 DF_4 num_conto \rightarrow nome_filiale



NON E' BCNF PERCHE' NESSUNA PARTE SINISTRA DELLE DFE' UNA SUPERCHIAVE

LA DECOMPOSIZIONE IN BCNF

R_1 (num_mutuo, valore)

R_2 (num mutuo, nome_filiale)

R_3 (num_conto, saldo)

R_4 (num mutuo, num conto, correntista)

PERDE LA DF_4 CIOE' LO STESSO num conto PUO' ESISTERE IN DUE FILIALI DIVERSE DANDO QUINDI PROBLEMI NELL'ESECUZIONE DEI JOIN

© Fabio A. Schreiber

Progettazione 34

ESEMPIO 4

RESTRIZIONE DI \mathcal{F}' DI \mathcal{F} ALLE R_1, \dots, R_4

$\mathcal{F}' = \{ \text{num_mutuo} \rightarrow \text{valore}, \text{num_mutuo} \rightarrow \text{nome_filiale}, \text{num_conto} \rightarrow \text{saldo} \}$

IN R_4 VALGONO SOLO DF BANALI

DF_4 num_conto \rightarrow nome_filiale $\in \mathcal{F}^+$ MA $\notin \mathcal{F}'^+$

E QUINDI

$$\mathcal{F}^+ \neq \mathcal{F}'^+$$

E LA DECOMPOSIZIONE NON E' CONSERVATIVA

© Fabio A. Schreiber

Progettazione 35

ESEMPIO 5

CONSIDERIAMO LA RELAZIONE

CORSO (**S**tudente, **S**ezione, **D**ocente)

CON I SEGUENTI VINCOLI:

1. OGNI SEZIONE DI UN CORSO HA UN UNICO DOCENTE

$$St, S \rightarrow D$$

2. OGNI DOCENTE INSEGNA IN UNA SOLA SEZIONE

$$D \rightarrow S$$

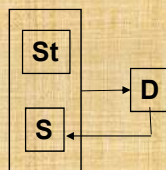
ESSA NON E' BCNF PERCHE' IN 2. D NON E' SUPERCHIAVE

© Fabio A. Schreiber

Progettazione 36

ESEMPIO 5

IL GRAFO DELLE DF PRESENTA UN CICLO CHE, PARTENDO DA UNA COPPIA, SI CHIUDE SU UN SOLO ATTRIBUTO



LA DECOMPOSIZIONE STANDARD FORNIREBBE

R_1 (Docente, Sezione)

R_2 (Studente, Docente)

PERDENDO LA $St, S \rightarrow D$ E QUINDI SAREBBE POSSIBILE INSERIRE TUPLE NELLE QUALI UNO STUDENTE SI TROVA AD AVERE DOCENTI DIVERSI PER LA STESSA MATERIA

© Fabio A. Schreiber

Progettazione 37

TERZA FORMA NORMALE: 3NF

UNO SCHEMA R E' IN 3NF SE, **PER TUTTE LE DF**
 $X \rightarrow Y$ ($X \subseteq R$, $Y \subseteq R$) CHE VALGONO SU R ,
VALE **ALMENO UNA** DELLE SEGUENTI
CONDIZIONI

- $X \rightarrow Y$ E' BANALE ($Y \subseteq X$)
- X E' UNA SUPERCHIAVE DI R
- Y E' CONTENUTA IN UNA CHIAVE CANDIDATA DI R

© Fabio A. Schreiber

Progettazione 38

TERZA FORMA NORMALE: 3NF

OGNI ATTRIBUTO NON APPARTENENTE ALLA CHIAVE

- DIPENDE FUNZIONALMENTE DA **OGNI** CHIAVE
- **NON DIPENDE TRANSITIVAMENTE** DA ALCUNA CHIAVE

- **BARCA**

(matr, nome, anno, modello, **lunghezza**, **stazza**, armatore)

NON E' IN 3NF PERCHE' `modello` NON FA PARTE DELLA
CHIAVE, MA VALE LA DIPENDENZA `modello` \rightarrow
`lunghezza`, `stazza`

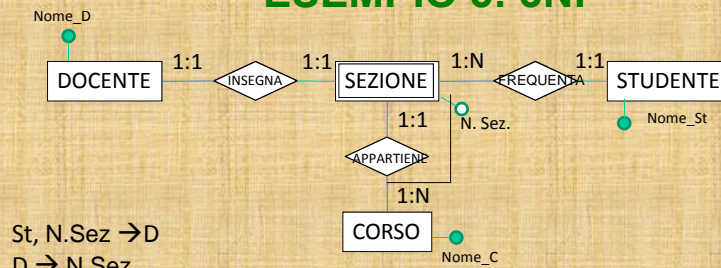
- **CORSO** (studente, **sezione**, **docente**)

E' IN 3NF POICHE' NELLA DF $D \rightarrow S$, S E' CONTENUTA IN S_t, S
CHE E' UNA CHIAVE CANDIDATA

© Fabio A. Schreiber

Progettazione 39

ESEMPIO 5: 3NF



DOCENTE (Nome D, ..., ~~N.Sez~~, ~~Nome C~~)
 CORSO (Nome C, ...)
 STUDENTE (Nome St, ...)
 SEZIONE (Nome C, N.sez., ...)
 FREQUENTA (Nome St, Nome C, N.sez., Nome D)

© Fabio A. Schreiber

Progettazione 40

ESEMPIO 5: 3NF

FREQUENTA (Nome St, Nome C, N.sez., Nome D)

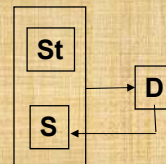
STUDENTE	SEZIONE	DOCENTE
Bianchi	Matematica A	Rossi
Bianchi	Basi di Dati A	Ceri
Morandi	Basi di Dati B	Tanca
Verdi	Matematica A	Rossi
Verdi	Fisica C	Bruni

INSEGNA E' CHIARAMENTE **RIDONDANTE**, MA EVITA LE ANOMALIE DI INSERIMENTO/ CANCELLAZIONE IN FREQUENTA

IL RISULTATO FINALE E' IMMEDIATAMENTE DEDUCIBILE DAL GRAFO DELLE DF

INSEGNA(Nome D, N.sez)

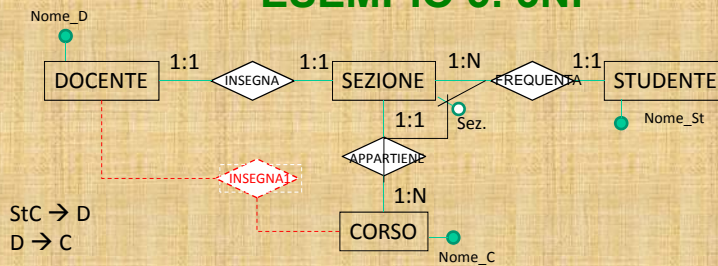
SEZIONE	DOCENTE
Matematica A	Rossi
Basi di Dati A	Ceri
Basi di Dati B	Tanca
Matematica A	Rossi
Fisica C	Bruni



© Fabio A. Schreiber

Progettazione 41

ESEMPIO 5: 3NF



DOCENTE (Nome_D, ...)

CORSO (Nome_C, ...)

STUDENTE (Nome_St, ...)

SEZIONE (Nome_St, Nome_C, sez., Nome_D)

INSEGNA1 (Nome_D, Nome_C)

STUDENTE	CORSO	DOCENTE
Bianchi	Matematica	Rossi
Bianchi	Basi di Dati	Ceri
Morandi	Basi di Dati	Tanca
Verdi	Matematica	Rossi
Verdi	Fisica	Bruni

© Fabio A. Schreiber

Progettazione 42

ALTRE FORME NORMALI: 1NF

- I DOMINI DEGLI ATTRIBUTI DEVONO COMPRENDERE SOLAMENTE VALORI **ATOMICI** (SEMPLICI ED INDIVISIBILI)
- OGNI VALORE DI UN ATTRIBUTO DEVE ESSERE **SINGOLO**
- ELIMINA LA POSSIBILITA' DI ANNIDARE RELAZIONI O TUPLE ENTRO ALTRE RELAZIONI
- **VIENE CONSIDERATA PARTE DELLA DEFINIZIONE STESSA DI RELAZIONE**

© Fabio A. Schreiber

Progettazione 43

ALTRE FORME NORMALI: 2NF

- PER **TUTTE** LE CHIAVI CANDIDATE, GLI ATTRIBUTI CHE NON APPARTENGONO ALLA CHIAVE **DEVONO DIPENDERE COMPLETAMENTE** DALLA CHIAVE
 - QUANDO LA CHIAVE E' FORMATA DA PIU' ATTRIBUTI E' **VIETATO** CHE CI SIANO DIPENDENZE TRA **SOLO ALCUNI ATTRIBUTI DELLA CHIAVE** E ALTRI ATTRIBUTI DELLO SCHEMA
- **MUTUO**
(numero, contraente, valore, filiale, patrimonio, città)
NON E' IN 2NF PERCHE' LA CHIAVE E' DATA DALLA COPPIA numero, filiale MA patrimonio, città DIPENDONO SOLO DA filiale E NON DA numero

© Fabio A. Schreiber

Progettazione 44

ALTRI ESEMPI

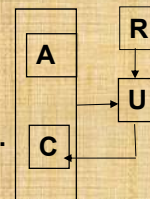
INFO_AEREI (Aeroporto, Compagnia, Ufficio, Responsabile)

DF₁: AC → U

DF₂: U → C

DF₃: R → U

E' IN **2NF** PERCHE' IN R → U NE' R NE' U SONO CHIAVE.
METTENDO IN CHIAVE R DIVENTA 3NF



COLTURE (Stagione, Pianta, Concime, Antiparass.)

DF₁: SP → C

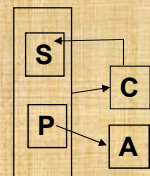
DF₂: C → S

DF₃: P → A

E' IN **1NF** PERCHE' P → A E' UNA DIPENDENZA PARZIALE

R₁(S, P, C)

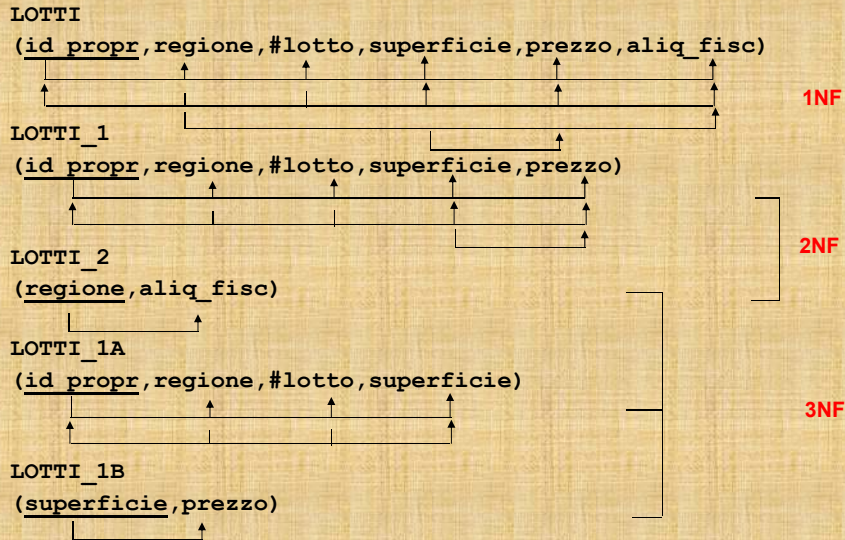
R₂(P, A)



© Fabio A. Schreiber

Progettazione 45

ESEMPIO DI NORMALIZZAZIONE



© Fabio A. Schreiber

Progettazione 46

ALTRE FORME NORMALI

- ESISTONO ALTRE FORME NORMALI CHE DERIVANO DA CONSIDERAZIONI TEORICHE, MA HANNO SCARSA IMPORTANZA PRATICA
- LA TEORIA DELLE DIPENDENZE FUNZIONALI E DELLA NORMALIZZAZIONE COSTITUISCE UN CAPITOLO IMPORTANTE DELLA TEORIA DELLE BASI DI DATI, PER LA QUALE SI RIMANDA ... AL PROSSIMO CICLO!

© Fabio A. Schreiber

Progettazione 47