

ESERCIZI DI BASI DI DATI

Raffaella Gentilini

INDICE

1. Algebra e Calcolo Relazionale	1
2. Progettazione Concettuale e Logica	5
3. Normalizzazione	9
3.1. Dipendenze Funzionali	9
3.2. Forme Normali	9
4. Soluzioni	10

1. ALGEBRA E CALCOLO RELAZIONALE

Esercizio 1. Sia dato il seguente schema relazionale:

- *persona* (*id_persona*, *nome*, *cognome*)
- *film* (*id_film*, *id_regista*, *titolo*, *genere*, *anno*): dove *id_regista* e' una chiave esterna che fa riferimento a *persona*;
- *partecipazione* (*id_attore*, *id_film*, *ruolo*): dove *id_attore* ed *id_film* sono chiavi esterne che fanno riferimento rispettivamente a *persona* e *film*;
- *cinema* (*id_cinema*, *nome*, *indirizzo*)
- *proiezione* (*id_cinema*, *id_film*, *giorno*): dove *id_cinema* e *id_film* sono chiavi esterne che fanno riferimento rispettivamente a *cinema* e *film*.

Con riferimento allo stato della base di dati riportato di seguito, si illustri il risultato delle interrogazioni in algebra relazionale 1.1–1.3.

	id_cinema	nome	indirizzo
• relazione <i>cinema</i>	02	S. Angelo	Via Lucida 6 Perugia
	01	Zenith	Via Bonfigli 11 Perugia
	03	Multisala Clarici	Corso Cavour 84 Foligno
	04	Multiplex Giometti	Strada Centova Perugia

	id_persona	nome	cognome
	01	Kidman	Nicole
	02	Bettany	Paul
	03	Watson	Emily
	04	Skarsgard	Stellan
	05	Travolta	John
	06	L. Jackson	Samuel
	07	Willis	Bruce
	08	Irons	Jeremy
	09	Spader	James
	10	Hunter	Holly
• relazione <i>persona</i>	11	Arquette	Rosanna
	12	Wayne	John
	13	von Trier	Lars
	14	Tarantino	Quentin
	15	Cronenberg	David
	16	Mazursky	Paul
	17	Jones	Grace
	18	Glen	John
	19	Eastwood	Clint
	20	Spacey	Kevin
	21	Mendes	Sam
	22	Jolie	Angelina

	id_attore	id_film	ruolo
	01	05	Grace
	02	05	Tom Edison
	03	04	Bess
	04	04	Jan
	05	03	Vincent Vega
	06	03	Jules Winnfield
	07	03	Butch Coolidge
• relazione <i>partecipazione</i>	08	02	Beverly
	09	01	James Ballard
	10	01	Helen Remington
	11	01	Gabrielle
	04	05	Chuck
	16	07	May Day
	19	08	John Wilson
	20	09	Jim Williams
	20	10	Lester Burnham

	id_cinema	id_film	giorno			
	02	05	01/05/2002			
	02	05	02/05/2002			
	02	05	03/05/2002			
	02	04	02/12/1996			
	01	01	07/05/1996			
	02	07	09/05/1985			
	01	04	02/08/1996			
	04	03	08/04/1994			
	03	06	02/12/1990			
	02	02	08/12/1990			
	03	03	06/11/1994			
• relazione <i>proiezione</i>	04	03	06/11/1994			
	01	06	05/07/1980			
	02	04	02/09/1996			
	04	06	01/08/2002			
	03	06	09/11/1960			
	01	02	12/03/1988			
	02	08	01/02/1989			
	02	01	11/05/1997			
	02	03	05/07/1994			
	02	06	01/08/2002			
	01	03	02/03/1994			
	02	09	02/12/2008			
	02	10	03/10/2000			
	02	11	02/03/2004			
	id_film	id_regista	titolo	genere	anno	
	01	15	Crash	drammatico	1996	
	02	15	Faux-Semblants	commedia	1988	
	03	14	Pulp Fiction	poliziesco	1994	
• relazione <i>film</i>	04	13	Breaking the Waves	drammatico	1996	
	05	13	Dogville	drammatico	2002	
	06	12	Alamo	western	1960	
	07	18	Pericolosamente tua	spionaggio	1985	
	08	19	Chasseur blanc, coeur noir	drammatico	1989	
	10	21	American Beauty	drammatico	1999	

- 1.1 $(\Pi_{titolo, genere} \sigma_{anno \leq 1988} \text{FILM}) \times \text{PERSONA}$
- 1.2 $\Pi_{titolo, nome, cognome}(\text{FILM} \bowtie_{id_regista=id_persona} \text{PERSONA})$
- 1.3 $(\Pi_{nome, cognome, titolo}(\text{FILM} \bowtie \text{PROIEZIONE} \bowtie_{id_attore=id_persona} \text{PERSONA})) \div$
 $\Pi_{titolo}(\text{FILM} \bowtie_{id_regista=id_persona} (\sigma_{cognome=Lars} \text{PERSONA}))$

Esercizio 2. Con riferimento allo schema relazionale dell'esercizio precedente, si esprimano in algebra relazionale le seguenti interrogazioni:

- 2.1 Elencare i titoli dei film drammatici mantenuti nella base di dati.
- 2.2 Definire la lista dei film proiettati al cinema Zenith.
- 2.3 Elencare nome e cognome di ogni regista presente nella base di dati.
- 2.4 Quali sono i nominativi (nome e cognome) degli attori?
- 2.5 Quali sono i nominativi (nome e cognome) degli artisti che sono tanto attori quanto registi?
- 2.6 Quali film sono stati proiettati prima del 2002?
- 2.7 Elencare i film diretti da Lars von Trier.
- 2.8 Definire la lista dei registi che hanno diretti sia film drammatici che commedie.
- 2.9 Elencare i film interpretati da Nicole Kidman e proiettati al cinema Zenith.
- 2.10 Elencare gli attori che non hanno mai interpretato ruoli in film drammatici.
- 2.11 Elencare gli attori che hanno interpretato qualche film proiettato al cinema Zenith dopo l'anno 2002.
- 2.12 Elencare i cinema che hanno proiettato tutti i film.
- 2.13 Elencare gli attori apparsi in ogni cinema mantenuto nella base di dati.
- 2.14 Elencare i registi piu' produttivi, ovvero i registi che hanno diretto il maggior numero di film.

Esercizio 3. Sia dato il seguente schema relazionale:

- *fornitore* (*id_for*, *Fnome*, *reddito*, *citta*)
- *componente* (*id_comp*, *Cnome*, *colore*, *peso*, *citta*)
- *progetto* (*id_prog*, *Pnome*, *citta*)
- *fornisce* (*id_for*, *id_comp*, *id_prog*, *quantita*)

Definiamo *originale* un prodotto fornito da un solo fornitore. Esprimere in algebra relazionale le seguenti interrogazioni:

- 3.1 Determinare i nomi dei fornitori che non forniscono alcun componente originale.
- 3.2 Determinare i nomi dei fornitori che forniscono solo componenti originali.
- 3.3 Determinare i nomi dei fornitori che forniscono al piu' un componente non originale.
- 3.4 Determinare i nomi dei fornitori che forniscono almeno un componente originale.
- 3.5 Determinare le coppie di nomi di fornitori tali che l'intersezione dei componenti da loro forniti sia vuota.
- 3.6 Determinare le citta' in cui risiedono almeno due fornitori con reddito maggiore o uguale a 100, escludendo le citta' cui non e' associato alcun progetto.
- 3.7 Per ogni citta, si determini il massimo ed il minimo dei pesi dei componenti ad esse associati.
- 3.8 Si determinino le citta' cui sono associati due o piu' componenti, ma non piu' di un progetto.

- 3.9 Si determinino i componenti di peso massimo.
- 3.10 Si determinino i componenti di peso minimo.
- 3.11 Si determini il reddito dei fornitori che gestiscono tutti i progetti cui viene fornito (da loro a da altri fornitori) almeno un componente di colore rosso.

Esercizio 4. Si riconsideri la soluzione dell'esercizio precedente, nell'ipotesi di voler utilizzare l'operatore di divisione e le funzioni aggregate solo se strettamente necessari (i.e di voler esprimere le interrogazioni 3:13:11 nell'algebra relazionale di base).

Esercizio 5. Si consideri lo schema di basi di dati che contiene le seguenti relazioni:

- *deputati* (codice, cognome, nome, commissione, provincia, collegio)
- *collegi* (provincia, numero, nome)
- *province* (sigla, nome, regione)
- *regioni* (codice, nome)
- *commissioni* (numero, nome, presidente)

Formulare in algebra relazionale e nel calcolo relazionale sulle tuple le seguenti interrogazioni:

- 5.1 Trovare nome e cognome dei presidenti di commissioni cui partecipa almeno un deputato eletto in unaprovincia della Sicilia
- 5.2 Trovare nome e cognome dei deputati della commissione bilancio.
- 5.3 Trovare nome, cognome e provincia di elezione dei deputati della commissione di bilancio
- 5.4 Trovare nome, cognome, provincia e regione di elezione dei deputati della commissione di bilancio
- 5.5 Trovare le regioni in cui vi sia un singolo collegio, indicando il nome ed il cognome del deputato ivi eletto.
- 5.6 Trovare i collegi della stessa regione in cui siano stati eletti deputati con lo stesso nome proprio.

2. PROGETTAZIONE CONCETTUALE E LOGICA

Esercizio 6. Si consideri una base di dati per la gestione delle partite di un campionato di calcio. Lo schema relazionale relativo ad una porzione della suddetta BD e' il seguente:

- SQUADRA(citta', sponsor, allenatore)
- GIORNATA_CAMPIONATO(numero, data)
- PARTITA(giornata, num, squadra_casa, squadra_ospite, goal1, goal2), dove l'attributo *squadra_casa* e' chiave esterna sulla relazione *squadra*, l'attributo *squadra_ospite* e' chiave esterna sulla relazione *squadra*, e l'attributo *giornata* e' chiave esterna sulla relazione *giornata_campionato*.

Nello schema relazionale delineato, l'attributo *num* nella relazione PARTITA si riferisce al numero progressivo della partita nella giornata di campionato indicata dall'attributo *giornata*. L'attributo *goal1* (rispettivamente, *goal2*) nella relazione PARTITA si riferisce al numero di goal segnati dalla squadra che gioca in casa (rispettivamente, ospite).

- (6.1) Si ricavi un diagramma EER corrispondente allo schema relazionale delineato nell'esercizio 1.
- (6.2) Si supponga che i requisiti per la costruzione della BD relativa alla gestione del campionato di calcio siano modificati/ampliati come segue:

- Per ogni squadra, si vogliono conoscere i giocatori tesserati per tale squadra con nome, cognome, data di nascita e ruolo di tesseramento.
- Per ogni partita, si vuole conoscere:
 - i giocatori che hanno giocato in tale partita ed i ruoli di tali giocatori (che possono cambiare di partita in partita)
 - nome, cognome, città e data di nascita dell'arbitro della partita
- Si vuole distinguere le partite giocate regolarmente dalle partite rinviate. In particolare, per le partite rinviate si vuole tenere traccia della data di effettivo svolgimento. Si vuole infine distinguere anche le partite giocate in città diverse da quelle delle città ospitanti, tenendo traccia del motivo dello spostamento e del nome della nuova città, teatro del campo neutro.

Si modifichi lo schema definito al punto precedente in modo da tenere conto dei requisiti modificati appena delineati.

Esercizio 7. Si supponga di voler progettare un client di posta elettronica. Il programma deve poter gestire un numero arbitrario di *account*. Quando un utente crea un account, specifica un unico indirizzo di posta elettronica, il proprio nome, l'indirizzo del server di posta in entrata, l'indirizzo del server di posta in uscita, il protocollo di comunicazione ed una descrizione dell'account. Non possono esserci due account associati allo stesso indirizzo di posta elettronica.

L'utente o il sistema possono creare una o più cartelle (*mailbox*) per contenere i messaggi (alcune mailbox sono create automaticamente dal sistema al momento della creazione di un account). Le mailbox possono essere associate ad uno specifico account, ma ciò non è obbligatorio. Tali mailbox sono identificate da un nome e per ciascuna di esse bisogna conoscere il numero totale di messaggi che contiene ed il numero di messaggi non letti. L'utente può spostare i messaggi (di qualunque account) da una mailbox ad un'altra. Ad ogni istante, ciascun messaggio deve trovarsi sempre in una ed una sola mailbox.

Un messaggio è identificato mediante un numero univoco, ha un mittente, uno o più destinatari (a meno che non sia salvato come draft senza essere spedito, nel qual caso può non aver alcun destinatario associato), un oggetto (*subject*), un testo ed eventualmente uno o più allegati (*attachment*). Per ogni messaggio bisogna conoscere la data di spedizione e di arrivo, nonché le informazioni relative all'intestazione (*header*) del messaggio stesso. Inoltre, ciascun messaggio può essere 'letto' oppure 'non letto' ed eventualmente marcato come 'importante'. Il client deve mantenere l'indirizzo e possibilmente il nome di ciascun contatto (destinatario o mittente): Si noti che il titolare di un account è un contatto.

Gli attributi di un allegato includono il nome, il tipo, la dimensione in byte e l'allegato vero e proprio.

- (7.1) Si progetti un diagramma EER che descriva il contenuto informativo del sistema, risolvendo le ambiguità del testo, prestando particolare attenzione a non introdurre ridondanze e indicando con chiarezza le assunzioni fatte. Eventuali vincoli di integrità non rappresentabili nel diagramma possono essere espressi nel linguaggio naturale.
- (7.2) Si traduca nel modello relazionale il diagramma individuato al punto precedente.

Esercizio 8. Si consideri il seguente schema di basi di dati, relativo ad una porzione di un DBMS per la gestione di un registro automobilistico:

- VEICOLO(targa, cilindrata, num_posti, tasse)
- PROPRIETARIO(nome, targa)

(8.1) Si ricavi un diagramma EER corrispondente allo schema relazionale delineato nell'esercizio 1.

(8.2) Si supponga che i requisiti per la costruzione della base di dati relativa alla gestione di un registro automobilistico siano modificati/ampliati come segue:

- di ciascun veicolo interessa registrare la targa, la cilindrata, i cavalli fiscali, la velocità, il numero di posti e la data di immatricolazione;
- i veicoli sono classificati in categorie (automobili, ciclomotori, camion ...)
- ciascun veicolo appartiene ad uno specifico modello; tra i dati relativi ai veicoli, vi è la codifica del tipo di combustibile utilizzato;
- di ciascun modello di veicolo è registrata la fabbrica di produzione ed il numero di versioni prodotte;
- ciascun veicolo può avere uno o più proprietari, che si succedono nel corso della vita del veicolo. Di ciascun proprietario interessa registrare nome, cognome, indirizzo e codice fiscale.

Si modifichi lo schema definito al punto precedente in modo da tenere conto dei requisiti modificati delineati sopra.

Esercizio 9. Si consideri una base di dati per la gestione di un concorso di musica, al quale partecipano diversi gruppi musicali che eseguono differenti generi di musica. Ciascun gruppo formato da un certo numero di musicisti (che devono essere maggiorenni), cui assegnato un identificatore numerico positivo, unico all'interno del gruppo (ma che può coincidere per musicisti di gruppi diversi). Lo schema relazionale utilizzato il seguente:

- STRUMENTO(nome, tipo)
- GRUPPO(nome, genere)
- MUSICISTA(id, gruppo, eta, strumento), dove l'attributo *gruppo* è chiave esterna sulla relazione *gruppo* e l'attributo *strumento* è chiave esterna sulla relazione *strumento*.

(9.1) Si ricavi un diagramma EER corrispondente allo schema relazionale delineato sopra.

(9.2) Si supponga che, in generale, un musicista possa suonare più strumenti e possa far parte di più gruppi. Si supponga inoltre che, invece di un codice numerico univoco all'interno di un gruppo, ciascun musicista sia univocamente identificato dal proprio codice fiscale. Si progetti un nuovo diagramma EER che tenga conto di questi requisiti e lo si traduca nel modello relazionale.

Esercizio 10. Si vuole progettare una BD di supporto alla gestione delle informazioni d'interesse per un amministratore di condomini, contenente le informazioni sotto specificate:

- Di un condominio, interessano l'indirizzo ed il numero del conto corrente dove vengono fatti i versamenti delle spese sostenute. Un condominio si

compone di un certo numero di appartamenti dei quali interessano il numero dell'interno, il numero dei vani, la superficie, lo stato (libero oppure occupato).

- Gli appartamenti possono essere locati. In tal caso, dell'inquilino interessano il nome, il codice fiscale, il telefono e il saldo, cioè la somma che l'inquilino deve all'amministratore condominiale per le spese sostenute. Alcuni appartamenti locati possono essere stati disdetti. In tal caso, interessa la data della disdetta.
- Un appartamento può avere più proprietari ed un proprietario può possedere più appartamenti. Di ogni proprietario interessano il nome, il codice fiscale, l'indirizzo, il telefono e la somma che il proprietario deve all'amministratore condominiale per le spese sostenute.
- Le spese riguardano i condomini e di esse interessano il codice di identificazione, la natura (luce, pulizia, ascensore, ...), la data e l'importo. Fra le spese si distinguono quelle straordinarie, a carico dei proprietari, ed ordinarie, a carico degli inquilini. Le spese ordinarie vengono pagate in un'unica rata, mentre le spese straordinarie possono essere pagate in più rate e di ognuna di esse occorre ricordare la data e l'importo.

Si definisca uno schema EER che descriva il contenuto informativo del sistema, illustrando con chiarezza le eventuali assunzioni fatte. Lo schema dovrà essere completato con attributi ragionevoli per ciascuna entità (identificando le possibili chiavi) e relazione. Vanno specificati accuratamente i vincoli di cardinalità e partecipazione di ciascuna relazione.

Esercizio 10. Si vuole progettare una base di dati di supporto alla gestione di un corso di laurea universitaria, contenente le seguenti informazioni:

- Per ogni docente, il nome, il dipartimento cui afferisce (ogni docente afferisce ad un unico dipartimento, ma docenti diversi, di uno stesso corso di laurea, possono afferire a dipartimenti diversi) il titolo (professore ordinario, associato o ricercatore), il numero dell'ufficio (che identifica univocamente l'ufficio all'interno del dipartimento di appartenenza del docente), il numero di telefono, l'indirizzo di posta elettronica.
- Per ogni membro del personale amministrativo, il nome, il tipo di lavoro svolto (segretario, tecnico di laboratorio, ...), il numero della stanza assegnatagli, appartenente all'insieme delle stanze a disposizione del corso di laurea (una stessa stanza può essere divisa da più membri del personale), un numero di telefono (uno distinto per ogni membro del personale).
- Per ogni corso, un codice identificativo, il nome del corso, il titolare (professore ordinario o associato), l'esercitatore (di tipo ricercatore, si assume che ogni corso abbia un unico esercitatore e che uno stesso ricercatore possa svolgere il ruolo di esercitatore per più corsi), il quadrimestre o i quadrimestri in cui il corso si tiene (si assuma un'organizzazione dell'anno accademico in 3 quadrimestri e una durata dei corsi che varia da 1 a 3 quadrimestri), giorni e fasce orarie delle lezioni (si assuma che, per i corsi distribuiti su più quadrimestri, giorni e fasce orarie rimangano invariati da un quadrimestre all'altro). aule (si assuma che non necessariamente le lezioni di un corso vengano tenute sempre nella stessa aula). Si definisca uno schema EER che descriva il contenuto informativo del sistema, illustrando

con chiarezza le eventuali assunzioni fatte. Lo schema dovrà essere completato con attributi ragionevoli per ciascuna entità (identificando le possibili chiavi) e relazione. Vanno specificati accuratamente i vincoli di cardinalità e partecipazione di ciascuna relazione.

3. NORMALIZZAZIONE

3.1. Dipendenze Funzionali.

Esercizio 11. Sia $R(A, B, C, D)$ uno schema di relazione. Si dimostri la correttezza o falsità delle seguenti regole di inferenza per dipendenze funzionali.

- (1) $\{AB \rightarrow D\} \models ABC \rightarrow D$
- (2) $\{ABC \rightarrow D\} \models AB \rightarrow D$
- (3) $\{C \rightarrow B, A \rightarrow D\} \models AC \rightarrow B$
- (4) $\{AB \rightarrow D, B \rightarrow C\} \models AC \rightarrow D$
- (5) $\{A \rightarrow B, AB \rightarrow D\} \models A \rightarrow D$

Esercizio 12. Sia $R(X)$ una relazione sull'insieme di attributi X e siano $Y \subseteq X, W \subseteq X, Z \subseteq X, K \subseteq X$. È vero o falso che se $Y \rightarrow W$ e $Z \rightarrow K$, allora $Y \cup (Z \setminus W) \rightarrow WK$? In caso di risposta affermativa, fornire una derivazione di $Y \cup (Z \setminus W) \rightarrow WK$ da $Y \rightarrow W$ e $Z \rightarrow K$ usando gli assiomi di Armstrong. Altrimenti, fornire un controesempio.

Esercizio 13. Siano dati lo schema di relazione $R(A, B, C, D, E, F, G, H, I, J)$ ed il relativo insieme di dipendenze funzionali $F = \{ABD \rightarrow E, AB \rightarrow G, B \rightarrow F, C \rightarrow J, CJ \rightarrow I, G \rightarrow H\}$.

- (8.1) Stabilire se F è o meno una copertura minimale. In caso di risposta negativa, determinare una copertura minimale di F .
- (8.2) Determinare l'insieme delle chiavi candidate di R .

Esercizio 14. Siano dati lo schema relazionale $R(A, B, C, D, E, F)$ e gli insiemi di dipendenza funzionali $G = \{AB \rightarrow C, B \rightarrow A, AD \rightarrow E, BD \rightarrow F\}$ ed $H = \{AB \rightarrow C, B \rightarrow A, AD \rightarrow EF\}$

- (9.1) Determinare una copertura minimale per G ed una copertura minimale per H .
- (9.2) Stabilire se G ed H sono equivalenti.

3.2. Forme Normali.

Esercizio 15. Si considerino lo schema di relazione $R(A, B, C, D, E, F)$ e l'insieme di dipendenze associato: $G = \{A \rightarrow B, C \rightarrow AD, AF \rightarrow EC\}$.

- (10.1) Si determinino le chiavi candidate di R .
- (10.2) Si stabilisca se R è in 3NF. Qualora non lo sia, si definisca una decomposizione di R in 3NF che conservi le dipendenze date.

Esercizio 16. Si considerino lo schema di relazione $R(A, B, C, D, E, G)$ e l'insieme di dipendenze associato: $F = \{E \rightarrow D, C \rightarrow B, CE \rightarrow G, B \rightarrow A\}$.

- (11.1) Si stabilisca se R è in 3NF. Qualora non lo sia, si definisca una decomposizione di R in 3NF che conservi le dipendenze date.
- (11.2) Se R non è in BCNF, determinare una decomposizione losslessjoin di R in BCNF.

4. SOLUZIONI

Esercizio 3: Soluzione. Nelle soluzioni proposte, gli operatori di divisione e funzioni aggregate verranno utilizzati solo se strettamente necessari.

- 3.1 Determinare i nomi dei fornitori che non forniscono alcun componente originale.

```
for ← Πid_for, id_comp(fornisce)
for1(id_for1, id_comp1) ← for
//componenti non originali
comp_non_orig ← Πid_comp(for ⋈id_for≠id_for1∧id_comp=id_comp1 for1)
T1 ← comp_non_orig × (Πid_for(for))
fornitori che forniscono almeno un componente originale
Bad ← Πid_for(for) – T1
//fornitori che non forniscono alcun componente originale
R ← ΠFnome(Fornitore ⋈ (Πid_for(for) – Bad))
```

- 3.2 Determinare i nomi dei fornitori che forniscono solo componenti originali.

```
for ← Πid_for, id_comp(fornisce)
for1(id_for1, id_comp1) ← for
//fornitori di almeno un componente non originale
Bad ← Πid_for(for ⋈id_for≠id_for1∧id_comp=id_comp1 for1)
//fornitori che forniscono solo componenti originali
Ris ← ΠFnome(Fornitore ⋈ (Πid_for(forn) – bad))
```

- 3.3 Determinare i nomi dei fornitori che forniscono al piu' un componente non originale.

```
for ← Πid_for, id_comp(fornisce)
for1(id_for1, id_comp1) ← for
//fornitori di almeno un componente non originale
T ← Πid_for, id_comp(for ⋈id_for≠id_for1∧id_comp=id_comp1 for1)
T1(id_for1, id_comp1) ← T
//fornitori di almeno due componenti non originali
Bad ← Πid_for(T ⋈id_for≠id_for1∧id_comp=id_comp1 T1)
//nomi dei fornitori che forniscono al piu' un componente non originale
R ← ΠFnome(Fornitore ⋈ (Πid_for(forn) – bad))
```

- 3.4 Determinare i nomi dei fornitori che forniscono almeno un componente originale.

```
for ← Πid_for, id_comp(fornisce)
for1(id_for1, id_comp1) ← for
// componenti non soggetti a copyright
T ← Πid_comp(for ⋈id_for≠id_for1∧id_comp=id_comp1 for1)
// componenti soggetti a copyright ovvero originali
T1 ← Πid_comp(componente – T)
//fornitori che forniscono almeno un componente originale
R ← ΠFnome(fornitore ⋈ (Πid_for(fornisce ⋈ T1)))
```

- 3.5 Determinare le coppie di fornitori tali che l'intersezione dei componenti da loro forniti sia vuota.

```
for ← Πid_for, id_comp(fornisce)
for1(id_for1, id_comp1) ← for
//coppie di fornitori che forniscono almeno un componente uguale
Bad ← Πid_for, id_for1(for ⋈id_for≠id_for1∧id_comp=id_comp1)
```

```

idfor ← Πid_for(fornitore)
idfor1(id_for1) ← for
coppie_id_for ← for × for1
R ← coppie_id_for – Bad

```

- 3.6 Determinare le citta' in cui risiedono almeno due fornitori con reddito maggiore o uguale a 100, escludendo le citta cui non e' associato alcun progetto.

```

T ← Πid_for,citta(σreddito ≥ 100(fornitore))
T1(id_for1, citta) ← T
R ← Πcitta(progetto) ⋈ Πcitta(T ⋈id_for ≠ id_for1 ∧ citta = citta1 T1)

```

- 3.7 Per ogni citta, si determini il massimo ed il minimo dei pesi dei componenti ad esse associati.

```

T(c, p) ← Πcitta,peso(componente)
T1(c1, p1) ← ttT
BadMax ← Πc,p(T ⋈c=c1 ∧ p < p1 T1)
GoodMax ← (c, pmax) ← cp – BadMax
BadMin ← Πc,p(T ⋈c=c1 ∧ p > p1 T1)
GoodMin ← (c, pmax) ← cp – BadMin
R ← GoodMin ⋈ GoodMax

```

- 3.8 Si determinino le citta' cui sono associati due o piu' componenti, ma non piu' di un progetto.

```

T(id_comp, c) ← Πid_comp,citta(componente)
T1(id_comp1, c1) ← T
2comp ← Πc(T ⋈c=c1 ∧ id_comp ≠ id_comp1 T1)
P(p, c) ← Πid_prog,citta(progetto)
P1(p1, c1) ← P
Bad ← Πc(P ⋈c=c1 ∧ p ≠ p1 P1)
R ← 2comp – Bad

```

- 3.9 Si determinino i componenti di peso massimo.

```

T(id, p) ← Πid,componente,peso(componente)
T1(id1, p1) ← T
BadMax ← Πid,p(T ⋈p < p1 T1)
GoodMax ← T – BadMax

```

- 3.10 Si determinino i componenti di peso minimo.

```

T(id, p) ← Πid,componente,peso(componente)
T1(id1, p1) ← T
BadMin ← Πid,p(T ⋈p > p1 T1)
GoodMin ← T – BadMin

```

- 3.11 Si determini il reddito dei fornitori che gestiscono tutti i progetti cui viene fornito (da loro a da altri fornitori) almeno un componente di colore rosso.

```

prod_rossi ← Πid_comp(fornisce ⋈ (σcolore='rosso'(componente)))
for ← Πid_for,id_prod(fornisce)
tutte_coppie ← Πid_for(fornisce) × prod_rossi
Bad ← Πid_for(tutte_coppie – for)
Good ← Πid_for(fornisce) – Bad

```

$$R \leftarrow \Pi_{\text{id,for,reddito}}(\text{Good} \bowtie \text{fornitore})$$

Esercizio 4: Soluzione. : Si consideri la soluzione dell'esercizio 3, in cui l'operatore di divisione e le funzioni aggregate sono state usate solo se strettamente necessari.

Esercizio 11: Soluzione. Sia $R(A, B, C, D)$ uno schema di relazione. Si dimostri la correttezza o falsità delle seguenti regole di inferenza per dipendenze funzionali.

(1) $\{AB \rightarrow D\} \models ABC \rightarrow D$

• Soluzione:

- (a) $AB \rightarrow D$ ipotesi
- (b) $ABC \rightarrow AB$ riflessività
- (c) $ABC \rightarrow D$ transitività da (a) e (b).

(2) $\{ABC \rightarrow D\} \models AB \rightarrow D$

• Soluzione: $\{ABC \rightarrow D\} \models AB \rightarrow D$ è falsa. Un controesempio è dato dalla seguente istanza r di R , dove $ABC \rightarrow D$ ma $AB \not\rightarrow D$:

A	B	C	D
a	b	c ₁	d ₁
a	b	c ₂	d ₂

(3) $\{C \rightarrow B, A \rightarrow D\} \models AC \rightarrow B$

• Soluzione:

- (a) $C \rightarrow B$ ipotesi
- (b) $A \rightarrow D$ ipotesi
- (c) $AC \rightarrow AB$ arricchimento (a) usando l'attributo A
- (d) $AB \rightarrow B$ riflessività
- (e) $AC \rightarrow B$ transitività da (c) e (d).

(4) $\{AB \rightarrow D, B \rightarrow C\} \models AC \rightarrow D$

• Soluzione: $\{AB \rightarrow D, B \rightarrow C\} \models AC \rightarrow D$ è falsa. Un controesempio è dato dalla seguente istanza r di R , dove $AB \rightarrow D$, $B \rightarrow C$, ma $AC \not\rightarrow D$:

A	B	C	D
a	a	a	a
a	b	a	b

(5) $\{A \rightarrow B, AB \rightarrow D\} \models A \rightarrow D$

• Soluzione:

- (a) $A \rightarrow B$ ipotesi
- (b) $AB \rightarrow D$ ipotesi
- (c) $A \rightarrow AB$ arricchimento (a) usando l'attributo A
- (d) $A \rightarrow D$ transitività da (b) e (c)

Esercizio 12: Soluzione. Sia $R(X)$ una relazione sull'insieme di attributi X e siano $Y \subseteq X, W \subseteq X, Z \subseteq X, K \subseteq X$. È vero o falso che se $Y \rightarrow W$ e $Z \rightarrow K$, allora $Y \cup (Z \setminus W) \rightarrow WK$? In caso di risposta affermativa, fornire una derivazione di $Y \cup (Z \setminus W) \rightarrow WK$ da $Y \rightarrow W$ e $Z \rightarrow K$ usando gli assiomi di Armstrong. Altrimenti, fornire un controesempio.

• Soluzione. La regola di inferenza è vera. Siano $Z = MN, W = MP$, con $N \cap P = \emptyset$. Allora:

- (a) $Y \rightarrow MP$ ipotesi

- (b) $MN \rightarrow K$ ipotesi
- (c) $YN \rightarrow MPN$ aumento (a) con N
- (d) $MNP \rightarrow KP$ aumento (b) con P
- (e) $MNP \rightarrow MKP$ aumento (d) con M
- (f) $YN \rightarrow MKP$ transitività da (c) ed (e).

Esercizio 13: Soluzione. Siano dati lo schema di relazione $R(A, B, C, D, E, F, G, H, I, J)$ ed il relativo insieme di dipendenze funzionali $F = \{ABD \rightarrow E, AB \rightarrow G, B \rightarrow F, C \rightarrow J, CJ \rightarrow I, G \rightarrow H\}$.

(8.1) Stabilire se F è o meno una copertura minimale. In caso di risposta negativa, determinare una copertura minimale di F .

- Soluzione. Applichiamo l'algoritmo per ottenere una copertura minimale.
 - Passo 1. I membri destri delle DF in F sono già unitari. Il passo 1 dell'algoritmo lascia dunque inalterato F .
 - Passo 2. Rimuoviamo dalle dipendenze gli attributi ridondanti. In $ABD \rightarrow E$ l'attributo A è ridondante sse $E \in BD^+$. Poiché $BD^+ = \{BDF\}$, concludiamo che A non è ridondante in $ABD \rightarrow E$
 L'attributo B è ridondante in $ABD \rightarrow E$ sse $E \in AD^+$. $AD^+ = \{AD\} \Rightarrow B$ non è ridondante in $ABD \rightarrow E$.
 L'attributo D è ridondante in $ABD \rightarrow E$ sse $E \in AB^+$. $AB^+ = \{ABFGH\} \Rightarrow D$ non è ridondante in $ABD \rightarrow E$.
 L'attributo A è ridondante in $AB \rightarrow G$ sse $G \in B^+$. $B^+ = \{BF\} \Rightarrow A$ non è ridondante in $AB \rightarrow G$.
 L'attributo B è ridondante in $AB \rightarrow G$ sse $G \in A^+$. $A^+ = \{A\} \Rightarrow B$ non è ridondante in $AB \rightarrow G$.
 L'attributo C è ridondante in $CJ \rightarrow I$ sse $I \in J^+$. $J^+ = \{J\} \Rightarrow C$ non è ridondante in $CJ \rightarrow I$.
 L'attributo J è ridondante in $CJ \rightarrow I$ sse $I \in C^+$. $C^+ = \{CIJ\} \Rightarrow C$ è ridondante in $CJ \rightarrow I$. Sostituiamo dunque $CJ \rightarrow I$ con $C \rightarrow I$, ottenendo $F = \{ABD \rightarrow E, AB \rightarrow G, B \rightarrow F, C \rightarrow J, C \rightarrow I, G \rightarrow H\}$
 - Passo 3. Eliminiamo infine le dipendenze ridondanti dall'insieme ottenuto al passo precedente. Per ogni dipendenza $X \rightarrow Y$ è sufficiente verificare se y appartiene alla chiusura di X rispetto ad $F \setminus \{X \rightarrow Y\}$

$X \rightarrow Y$	X^+ rispetto a $F \setminus \{X \rightarrow Y\}$	$X \rightarrow Y$ è ridondante?
$ABD \rightarrow E$	$ABD^+ = \{ABDGFH\}$	No
$AB \rightarrow G$	$AB^+ = \{ABF\}$	No
$B \rightarrow EF$	$B^+ = \{B\}$	No
$C \rightarrow J$	$C^+ = \{CI\}$	No
$C \rightarrow I$	$C^+ = \{CJ\}$	No
$G \rightarrow H$	$G^+ = \{G\}$	No

La copertura minimale richiesta e' dunque:

$$F = \{ABD \rightarrow E, AB \rightarrow G, B \rightarrow F, C \rightarrow J, C \rightarrow I, G \rightarrow H\}$$

(8.2) Determinare l'insieme delle chiavi candidate di R .

Gli attributi A, B, C, D devono far parte di ogni chiave poiche', non comparando a destra di alcuna DF in F , non possono essere derivati. Dunque, ogni chiave condidata K e' tale che $K \supseteq \{A, B, C, D\}$. Si ha $ABCD^+ = ABCDEFGHIJ = R$. Dunque, $ABCD$ e' una superchiave, rispetta il vincolo di minimalita' ed e' l'unica chiave condidata di R .

Esercizio 14: Soluzione. Siano dati lo schema relazionale $R(A, B, C, D, E, F)$ e gli insiemi di dipendenza funzionali $G = \{AB \rightarrow C, B \rightarrow A, AD \rightarrow E, BD \rightarrow F\}$ ed $H = \{AB \rightarrow C, B \rightarrow A, AD \rightarrow EF\}$

(9.1) Determinare una copertura minimale per G ed una copertura minimale per H .

• Soluzione. Calcoliamo una copertura minimale per G con l'algoritmo visto a lezione.

– Passo 1. I membri destri sono gia' unitari e dunque il primo passo non apporta modifiche a G .

– Passo 2. Rimuoviamo gli attributi ridondanti da ogni dipendenza.

L'attributo A e' ridondante in $AB \rightarrow C$ sse $C \in B^+$. $B^+ = \{BAC\} \supseteq \{B\}$. A e' dunque ridondante in $AB \rightarrow C$ che viene sostituita con $B \rightarrow C$.

L'attributo A e' ridondante in $AD \rightarrow E$ sse $E \in D^+$. $D^+ = \{D\} \Rightarrow A$ non e' ridondante in $AD \rightarrow E$.

L'attributo D e' ridondante in $AD \rightarrow E$ sse $E \in A^+$. $A^+ = \{A\} \Rightarrow D$ non e' ridondante in $AD \rightarrow E$.

L'attributo B e' ridondante in $BD \rightarrow F$ sse $F \in D^+$. $D^+ = \{D\} \Rightarrow B$ non e' ridondante in $BD \rightarrow F$.

L'attributo D e' ridondante in $BD \rightarrow F$ sse $F \in B^+$. $B^+ = \{B\} \Rightarrow D$ non e' ridondante in $BD \rightarrow F$.

– Passo 3. Eliminiamo infine le dipendenze ridondanti dall'insieme ottenuto al passo precedente. Per ogni dipendenza $X \rightarrow Y$ e' sufficiente verificare se y appartiene alla chiusura di X rispetto ad $F \setminus \{X \rightarrow Y\}$

$X \rightarrow Y$	X^+ rispetto a $F \setminus \{X \rightarrow Y\}$	$X \rightarrow Y$ e' ridondante?
$B \rightarrow C$	$B^+ = \{BA\}$	No
$B \rightarrow A$	$B^+ = \{BC\}$	No
$AD \rightarrow E$	$AD^+ = \{AD\}$	No
$BD \rightarrow F$	$BD^+ = \{BDCAE\}$	No

L'insieme di DF:

$$\{B \rightarrow C, B \rightarrow A, AD \rightarrow E, BD \rightarrow F\}$$

e' dunque una copertura per G . Operando analogamente su H otteniamo la seguente copertura minimale:

$$\{B \rightarrow C, B \rightarrow A, AD \rightarrow E, AD \rightarrow F\}$$

(9.2) Stabilire se G ed H sono equivalenti.

- Soluzione. Dobbiamo verificare che g e' coperto da H ed H e' coperto da G . Verifichiamo se G e' coperto da H ovvero $G \subseteq H^+$. Le dipendenze $AB \rightarrow C, B \rightarrow A, AD \rightarrow E$ in G appartengono anche ad H e dunque ad H^+ . Vediamo se $BD \rightarrow F \in H^+$. $BD \rightarrow F \in H^+$ sse $F \in BD^+$ (rispetto ad H). BD^+ rispetto ad H equivale a $\{BDACEF \supseteq \{F\}\}$. Possiamo dunque concludere che $G \subseteq H^+$. Al fine di provare $H \subseteq G^+$ dobbiamo verificare se $AD \rightarrow F \in G^+$. Si ha $F \notin AD^+$ (rispetto a G). Infatti $AD^+ = \{ADE\}$. Dunque $H \not\subseteq G^+$ e G ed H non sono equivalenti.

Esercizio 15: Soluzione. Si considerino lo schema di relazione $R(A, B, C, D, E, F)$ e l'insieme di dipendenze associato: $G = \{A \rightarrow B, C \rightarrow AD, AF \rightarrow EC\}$.

(10.1) Si determinino le chiavi candidate di R .

- Soluzione. Poiche' l'attributo F non compare nella parte destra di alcuna DF, ne segue che F deve appartenere ad ogni chiave candidata. Al contrario, $D, E,$ e B compaiono solo nella parte destra di DF. Ne segue che $D, E,$ e B non appartengono ad alcuna chiave candidata. Da:

$$\begin{aligned} - AF^+ &= AFBDEC = R \\ - CF^+ &= CFADEC = R \\ - F^+ &= F \end{aligned}$$

segue che AF e CF sono le uniche chiavi candidate di R .

(10.2) Si stabilisca se R e' in 3NF. Qualora non lo sia, si definisca una decomposizione di R in 3NF che conservi le dipendenze date.

- Soluzione. R non e' in 3NF. Si consideri infatti la DF $A \rightarrow B$. Tale dipendenza viola la 3NF in quanto A non e' una superchiave e B non e' primo. Calcoliamo una copertura minimale per G utilizzando l'algoritmo visto a lezione. Otteniamo:

$$G' = \{A \rightarrow B, C \rightarrow A, C \rightarrow D, AF \rightarrow E, AF \rightarrow C\}$$

da cui otteniamo la decomposizione: $R_1 = (AB), R_2 = (CAD), R_3 = (AFEC)$.

Esercizio 16: Soluzione. Si considerino lo schema di relazione $R(A, B, C, D, E, G)$ e l'insieme di dipendenze associato: $F = \{E \rightarrow D, C \rightarrow B, CE \rightarrow G, B \rightarrow A\}$.

(11.1) Si stabilisca se R e' in 3NF. Qualora non lo sia, si definisca una decomposizione di R in 3NF che conservi le dipendenze date.

- Soluzione. Determiniamo innanzitutto le chiavi della relazione. C ed E non compaiono a destra in alcuna DF, quindi devono appartenere ad ogni chiave. Si ha:

$$\begin{aligned} - CE^+ &= CEDBGA = R \\ - C^+ &= CBA \\ - E^+ &= ED \end{aligned}$$

Dunque CE e' l'unica chiave candidata di R .

R non e' in 3NF: La DF $E \rightarrow D$ e' tale che E non e' una superchiave e D non e' un attributo primo. Applichiamo dunque ad R l'algoritmo

visto a lezione per definire una decomposizione 3NF che preserva le dipendenze.

- Passo 1. Mediante l'algoritmo apposito, ci assicuriamo che F sia una copertura minimale (lo e').
 - Passo 2. Da $F = \{E \rightarrow D, C \rightarrow B, CE \rightarrow G, B \rightarrow A\}$ otteniamo la decomposizione: $R1 = (ED), R2 = (CB), R3 = (CEG), R4 = (BA)$
 - Passo 3. $R3$ contiene una chiave di R e dunque la decomposizione rimane invariata.
- (11.2) Se R non e' in BCNF, determinare una decomposizione losslessjoin di R in BCNF.
- Soluzione. La scomposizione e' in BCNF. Infatti:
 - DE,BC,AB sono relazioni binarie e dunque in BCNF
 - R3=CEG rispetta la BCNF poiche' per ogni $X \subseteq \{C, E, G\}$, si ha: X^+ contiene tutti gli attributi di $R3$ oppure X^+ non include attributi di $R3 \setminus X$.