



PEGASO
Università Telematica



Indice

1. INTRODUZIONE.....	3
2. VINCOLI DI INTEGRITÀ	4
3. CLASSIFICAZIONE DEI VINCOLI.....	6
4. VINCOLI DI TUPLA	8
5. CHIAVI.....	10
5.1 APPROFONDIMENTI	11
BIBLIOGRAFIA	12

1. Introduzione

In questa unità didattica si introducono, con esempi i seguenti concetti:

- Vincoli di integrità di un database;
- Vincoli di enunpla;
- Chiavi.

Ed inoltre vengono proposti esempi pratici sui quali ragionare.

2. Vincoli di integrità

Le strutture del modello relazionale ci permettono di organizzare le informazioni di interesse per le applicazioni di interesse. In molti casi però non è vero che qualsiasi insieme di tuple sullo schema rappresenti informazioni corrette per l'applicazione. In questo paragrafo si approfondisce anche con riferimento a relazioni prive di valori nulli (studiate nella UDA 7).

Esistono tuttavia istanze di basi di dati che, pur sintatticamente corrette, non rappresentano informazioni possibili per l'applicazione di interesse. Ad esempio, in Tabella 1, Tabella 2 e Tabella 3 si può notare una base di dati *scorretta*. Infatti si presentano le seguenti incongruità:

1. Nella prima tupla della relazione ESAMI, appare un voto pari a 36 che, nel sistema universitario italiano non è ammissibile, in quanto i voti devono essere compresi tra 0 e 30 o 30 e lode mentre per il superamento dell'esame tra 18 e 30;
2. Nella seconda tupla sempre della relazione *Esami*, appare il valore *28 e lode*, valore impossibile poiché la lode si può dare solo al valore di *Voto=30*;
3. Le ultime due tuple della relazione STUDENTI, mostrano due studenti diversi con lo stesso numero di matricola. Questo non risulta possibile poiché il numero di matricola dovrebbe individuare univocamente uno studente;
4. La quarta tupla della relazione ESAMI, presenta, per l'attributo STUDENTE, un valore dell'attributo matricola che non compare fra i numeri di matricola nella relazione STUDENTI: anche questa è una situazione indesiderabile, poiché i numeri di matricola ci forniscono informazioni solo come tramite verso le corrispondenti tuple della relazione STUDENTI.

Tabella 1: la relazione STUDENTI.

Matricola	Cognome	Nome	Data di nascita
200768	Verdi	Fabio	12/02/1997
937653	Rossi	Luca	10/10/1996
937653	Bruni	Mario	01/12/1996

Tabella 2: la relazione ESAMI.

Studente	Voto	Lode	Corso
200768	36		05
937653	28	Lode	01
937653	30	Lode	04
276545	25		01

Tabella 3: la relazione CORSI.

Codice	Titolo	Docente
01	Analisi	Giani
03	Chimica	Melli
04	Chimica	Belli

In una base di dati, situazioni come quella precedente non devono capitare poiché ne renderebbero la gestione molto problematica. Per questo vengono introdotti i *Vincoli di integrità* come proprietà che devono essere soddisfatte dalle istanze che rappresentano informazioni corrette per le applicazioni. Un vincolo può essere rappresentato come una funzione booleana (un predicato) che associa ad ogni istanza il valore *vero* o *falso*. Quindi le motivazioni per l'introduzione dei vincoli di integrità possono essere così riassunte:

- Una descrizione più accurata della realtà;
- Un contributo alla “qualità dei dati”;
- Utili nella progettazione (vedremo in seguito);
- Usati dai DBMS nella esecuzione delle interrogazioni;
- Alcuni tipi di vincoli (ma non tutti) sono “supportati” dai DBMS:
 - possiamo quindi specificare vincoli di tali tipi nella nostra base di dati e il DBMS ne impedisce la violazione.
- Per i vincoli “non supportati”, la responsabilità della verifica è dell'utente o del programmatore.

3. Classificazione dei vincoli

È possibile classificare i vincoli a seconda degli elementi di una base di dati che ne sono coinvolti. Si distinguono due categorie, la prima delle quali presenta alcuni casi particolari:

- vincoli *intrarelazionali*. Il soddisfacimento del vincolo è definito rispetto a singole relazioni della base di dati; i vincoli discussi precedentemente corrispondono a tale tipologia di vincolo. Talvolta il coinvolgimento riguarda le tuple (o addirittura i valori) separatamente le una dalle altre (Atzeni et al., 2017):
 - vincoli su valori (o di dominio). È il caso di un vincolo definito con riferimento a singoli valori (ad esempio il voto d'esame compreso tra 18 e 30). Impone una restrizione sul dominio dell'attributo;
 - vincoli di ennupla o di tupla. È un vincolo che può essere valutato su ciascuna tupla indipendentemente dalle altre. I vincoli relativi ai primi due casi rientrano in tale categoria.
- vincoli *interrelazionali*. Se coinvolge più relazioni; è questo il caso del quarto esempio in cui la situazione indesiderata può essere vietata richiedendo che un numero di matricola compaia nella relazione ESAMI solo se compare nella relazione STUDENTI.

Come esercizio, lo studente classifichi le violazioni dei vincoli di integrità delle relazioni riportate nelle tabelle sottostanti.

Codice fiscale	Cognome	Nome	Città di nascita	Pr	Data di nascita
LBRRSS82B09H501S	Rossi	Alberto	Roma	RM	10/02/1982
GSPVRD00C14F205D	Verdi	Giuseppe	Milano	MI	14/03/2000
FLPGLL99E14F839P	Gialli	Filippo	Napoli	NA	14/05/1999
RSSNNA80A5	Rossini	Anna	Firenze	FI	18/01/1980

Tabella 4: esempi di relazioni che rompono i vincoli di integrità. La tabella riportata sopra è la tabella CONTRIBUENTI mentre la tabella sottostante è la tabella TASSE. L'attributo Codice Fiscale funziona come chiave e come tale...

Codice fiscale	Anno	Imponibile	Tasse pagate
LBRSS82B09H501S	2018	20.000,00	12.000,00
GSPVRD00C14F205D	2018	35.000,00	15.000,00
GSPVRD00C14F205D	2018	45.000,00	20.000,00
FLPGLL99E14F839P	2018	60.000,00	25.000,00
RSTGTY88I17G7828H	2018	35.000,00	12.500,00

4. Vincoli di tupla

I vincoli di tupla esprimono condizioni sui valori di ciascuna tupla, indipendentemente dalle altre tuple. Come caso particolare si ha il vincolo di dominio che coinvolge un solo attributo. Una possibile sintassi è la seguente:

- espressione booleana di atomi che confrontano valori di attributo o espressioni aritmetiche su di essi. Ad esempio, nel caso della tabella ESAMI, precedentemente studiata, si potrebbero usare i seguenti vincoli:
 - $(Voto \geq 18) \text{ AND } (Voto \leq 30)$ per quanto riguarda il valore del voto;
 - $(Voto = 30) \text{ OR NOT } (Lode = "Lode")$ esprime il fatto che la lode deve essere abbinata solo al voto=30.

In Figura 1 un esempio di vincolo di tupla dove l'attributo *Lordo* dipende da altri due attributi *Ritenute* e *Netto*. Questo è un altro vincolo di tupla.

Vincoli di enupla,
altro esempio

Stipendi	Impiegato	Lordo	Ritenute	Netto
	Rossi	55.000	12.500	42.500
	Neri	45.000	10.000	35.000
	Bruni	47.000	11.000	36.000

$Lordo = (Ritenute + Netto)$

Figura 1: relazione STIPENDI.

In Figura 2 un esempio di violazione del suddetto vincolo di tupla.

Vincoli di enupla,
violazione

→

Stipendi	Impiegato	Lordo	Ritenute	Netto
	Rossi	55.000	12.500	42.500
	Neri	45.000	10.000	35.000
	Bruni	50.000	11.000	36.000

$Lordo = (Ritenute + Netto)$

Figura 2: esempio di violazione di vincolo di tupla.

Per quanto riguarda l'identificazione univoca delle ennuple, rappresentate nella relazione di Figura 3, possiamo osservare che:

- non ci sono due ennuple con lo stesso valore sull'attributo **Matricola**;
- non ci sono due ennuple uguali su tutti e tre gli attributi **Cognome, Nome e Data di Nascita**.

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/98
78763	Rossi	Mario	Ing Inf	3/11/96
65432	Neri	Piero	Ing Mecc	10/7/99
87654	Neri	Mario	Ing Inf	3/11/96
67653	Rossi	Piero	Ing Mecc	5/12/98

Figura 3: un esempio di identificazione delle ennuple. Relazione *STUDENTI*.

Queste osservazioni, di individuazione delle ennuple di una relazione, ci introducono al prossimo argomento che tratta proprio come rendere univoca l'individuazione delle ennuple di una relazione.

5. Chiavi

In questo paragrafo introduciamo i vincoli di chiave che sono i più importanti del modello relazionale. Nell'esempio di Figura 3, si nota che i valori delle varie tuple sull'attributo *Matricola* sono tutti diversi uno dall'altro: ciò consente di individuare univocamente gli studenti.

Possiamo a questo punto definire una **chiave** come un insieme di attributi che individua univocamente le tuple di una relazione. Formalizzando tale concetto:

- Un insieme K di attributi è **superchiave** per la relazione R se R **non contiene** due ennuple distinte t_1 e t_2 con $t_1[K] = t_2[K]$.
- K è chiave per R se è una superchiave minimale per R (cioè non contiene un'altra superchiave).

Nel caso della relazione STUDENTI di Figura 3, si ha:

- L'attributo Matricola è una chiave perchè:
 - è superchiave;
 - contiene un solo attributo e quindi è minimale.

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/98
78763	Rossi	Mario	Ing Inf	3/11/96
65432	Neri	Piero	Ing Mecc	10/7/99
87654	Neri	Mario	Ing Inf	3/11/96
67653	Rossi	Piero	Ing Mecc	5/12/98

Figura 4: relazione STUDENTI con le chiavi evidenziate.

- L'insieme {Cognome, Nome, Nascita}, opportunamente evidenziate in Figura 4 è una superchiave la quale è anche chiave;
- L'insieme {Matricola, Corso} è superchiave ma non è minimale poiché contiene un suo sottoinsieme proprio {Matricola} la quale è essa stessa superchiave minimale;
- L'insieme {Nome, Corso} non è superchiave perchè nella relazione compaiono due tuple, le prime due, fra loro uguali sia nel nome che nel corso.

5.1 Approfondimenti

Per fissare meglio il concetto di chiave, osserviamo la relazione rappresentata dalla Tabella 5 dove non sono contenute tuple tra loro uguali sia su **Cognome** sia su **Corso**. Ne consegue che l'insieme {**Cognome**, **Corso**} è superchiave. Poiché vi sono tuple uguali su Cognome e su Corso, tale insieme è superchiave minimale e cioè chiave.

Tabella 5: relazione di approfondimento per le chiavi.

Matricola	Cognome	Nome	Data di nascita	Corso
6328	Rossi	Dario	29/04/1994	Ing. Informatica
4766	Rossi	Luca	01/05/1995	Ing. Civile
4856	Neri	Luca	01/05/1995	Ing. Meccanica
5536	Neri	Luca	05/03/1993	Ing. Civile

Si rifletta sul fatto che tale chiave in questa relazione identifica tuple diverse ma ciò non sempre risulta vero: possono verificarsi omonimie nei cognomi di due studenti che seguono lo stesso corso. In questo caso si afferma che {Cognome, Corso} è casualmente una chiave per la relazione in questione.

Riassumendo, per i vincoli:

- i vincoli corrispondono a proprietà del mondo reale modellato dalla base di dati;
- interessano a livello di schema (con riferimento cioè a tutte le istanze);
- ad uno schema associamo un insieme di vincoli e consideriamo corrette (valide, ammissibili) le istanze che soddisfano tutti i vincoli;
- un'istanza può soddisfare altri vincoli ("per caso").

Inoltre:

- Una relazione non può contenere ennuple distinte ma uguali;
- Ogni relazione ha come superchiave l'insieme degli attributi su cui è definita e quindi ha (almeno) una chiave;
- L'esistenza delle chiavi garantisce l'accessibilità a ciascun dato della base di dati;
- Le chiavi permettono di correlare i dati in relazioni diverse:
 - Il modello relazionale è basato su valori.

Bibliografia

- Atzeni P., Ceri S., Fraternali P., Paraboschi S., Torlone R. (2018). Basi di Dati. McGraw-Hill Education.
- Batini C., Lenzerini M. (1988). Basi di Dati. In Cioffi G. and Falzone V. (Eds). Calderini. Seconda Edizione.