Capitolo 74

DBMS e SQL

74.1 Introduzione ai DBMS

«

Un DBMS (*Data base management system*) è, letteralmente, un sistema di gestione di ***basi di dati***, che per attuare questa gestione utilizza il software. Queste «basi di dati» sono dei contenitori atti a immagazzinare una grande quantità di dati, per i quali il «sistema di gestione» è necessario a permetterne la fruizione (diverso è il problema della semplice archiviazione dei dati).

74.1.1 Caratteristiche fondamentali

Un DBMS, per essere considerato tale, deve avere caratteristiche determinate. Le più importanti che permettono di comprenderne il significato sono elencate di seguito.

- Un DBMS è fatto per gestire grandi quantità di dati.

  Convenzionalmente si può intendere che un gruppo di informazioni sia di grandi dimensioni quando questo non possa essere contenuto tutto simultaneamente nella memoria centrale dell'elaboratore. In generale un DBMS non dovrebbe porre limiti alle dimensioni, tranne quelle imposte dai supporti fisici in cui devono essere memorizzate le informazioni.

- I dati devono poter essere condivisibili.

  L'idea che sta alla base dei sistemi di gestione dei dati è quella di accentrare le informazioni in un sistema di amministrazione unico. In tal senso è poi necessario che questi dati siano condivisibili da diverse applicazioni e da diversi utenti.

- I dati devono essere persistenti e affidabili.

  I dati sono persistenti quando continuano a esistere dopo lo spegnimento della macchina con cui vengono elaborati; sono affidabili quando gli eventi per cui si possono produrre alterazioni accidentali sono estremamente limitati.

- L'accesso ai dati deve essere controllabile.

  Dovendo trattare una grande mole di dati in modo condiviso, è indispensabile che esistano dei sistemi di controllo degli accessi, per evitare che determinate informazioni possano essere ottenute

da chi non è autorizzato, oppure che vengano modificate da chi non ne è il responsabile.

74.1.1.1 Livelli di astrazione dei dati

I dati gestiti da un DBMS devono essere organizzati a diversi livelli di astrazione. Generalmente si distinguono tre livelli: esterno, logico e interno.

- Il livello interno è quello usato effettivamente per la memorizzazione dei dati. In pratica è rappresentato dai file che contengono effettivamente le informazioni e dal modo con cui questi file vengono utilizzati. Il livello interno non è importante per la progettazione di una base di dati e nemmeno per la scrittura di programmi che devono interagire con il DBMS.
- Il livello logico, o concettuale, è quello che descrive i dati secondo la filosofia del DBMS particolare con cui si ha a che fare.
- Lo schema esterno è un'astrazione aggiuntiva che permette di definire dei punti di vista differenti dei dati descritti a livello logico. L'accesso ai dati avviene solo a questo livello, anche se di fatto può coincidere con il livello logico.

74.1.1.2 Ruoli e sigle standard

Lo studio sui DBMS ha generato degli acronimi che rappresentano persone o componenti essenziali di un sistema del genere. Questi acronimi devono essere conosciuti perché se ne fa uso abitualmente nella documentazione riferita ai DBMS.

L'organizzazione di una base di dati è compito del suo amministratore, definito DBA (*Data base administrator*). Eventualmente può trattarsi anche di più persone; in ogni caso, chi ha la responsabilità dell'amministrazione di uno o più basi di dati è un DBA.

Il concetto di DBA non è molto preciso, perché include assieme ruoli che possono essere differenti. Si va dall'amministratore dell'intero DBMS, fino a coloro che hanno la responsabilità di una base di dati singola.

La definizione della struttura dei dati (sia a livello logico, sia a livello esterno) viene fatta attraverso un linguaggio di programmazione definito DDL (*Data definition language*), la gestione dei dati avviene attraverso un altro linguaggio, detto DML (*Data manipulation language*), infine, la gestione della sicurezza viene fatta attraverso un linguaggio DCL (*Data control language*). Nella pratica, DDL, DML e DCL possono coincidere, come nel caso del linguaggio SQL.

74.1.2 Modello relazionale

Una base di dati può essere impostata secondo diversi tipi di modelli (logici) di rappresentazione. Quello più comune, che è anche il più semplice dal punto di vista umano, è il modello relazionale. In tal senso, un DBMS relazionale viene anche definito semplicemente come RDBMS.

Nel modello relazionale, i dati sono raccolti all'interno di ***relazioni***. Ogni relazione è una raccolta di nessuna o più ***tuple*** di tipo omogeneo. La tupla rappresenta una singola informazione completa, in rapporto alla relazione a cui appartiene, suddivisa in ***attributi***. Le in-

formazioni che possono essere inserite in ogni singolo attributo, dipendono da un ***dominio***. Una relazione, nella sua definizione, non ha una «forma» particolare, tuttavia questo concetto si presta a una rappresentazione tabellare: gli attributi sono rappresentati dalle colonne e le tuple dalle righe. Si osservi l'esempio della figura 74.1.

Figura 74.1. Relazione '**`Indirizzi (Cognome, Nome, Indirizzo, Telefono)`**'.

Indirizzi			
Cognome	**Nome**	**Indirizzo**	**Telefono**
Pallino	Pinco	Via Biglie 1	0222,222222
Tizi	Tizio	Via Tazi 5	0555,555555
Cai	Caio	Via Caini 1	0888,888888
Semproni	Sempronio	Via Sempi 7	0999,999999

In una relazione, le tuple non hanno una posizione particolare, sono semplicemente contenute nell'insieme della relazione stessa. Se l'ordine ha una rilevanza per le informazioni contenute, questo elemento dovrebbe essere aggiunto tra gli attributi, senza essere determinato da un'ipotetica collocazione fisica. Osservando l'esempio, si intende che l'ordine delle righe non ha importanza per le informazioni che si vogliono trarre; al massimo, un elenco ordinato può facilitare la lettura umana, quando si è alla ricerca di un nome particolare, ma ciò non ha rilevanza nella struttura che deve avere la relazione corrispondente.

Il fatto che la posizione delle tuple all'interno della relazione non

sia importante, significa che non è necessario poterle identificare: le tuple si distinguono in base al loro contenuto. In questo senso, una relazione non può contenere due tuple uguali: la presenza di doppioni non avrebbe alcun significato.

A differenza delle tuple, gli attributi devono essere identificati attraverso un nome. Infatti, il semplice contenuto delle tuple non è sufficiente a stabilire di quale attributo si tratti. Osservando la prima riga dell'esempio, diventa difficile distinguere quale sia il nome e quale il cognome:

Pallino	Pinco	Via Biglie 1	0222,222222

Assegnando agli attributi (cioè alle colonne) un nome, diventa indifferente la disposizione fisica di questi all'interno delle tuple.

74.1.2.1 Relazioni collegate

Generalmente, una relazione da sola non è sufficiente a rappresentare tutti i dati riferiti a un problema o a un interesse della vita reale. Quando una relazione contiene tante volte le stesse informazioni, è opportuno scinderla in due o più relazioni più piccole, collegate in qualche modo attraverso dei riferimenti. Si osservi il caso delle relazioni rappresentate dalle tabelle che si vedono nella figura 74.3.

Figura 74.3. Relazioni di un'ipotetica gestione del magazzino.

Articoli

Codice	Descrizione	Fornitore1	Fornitore2
vite30	Vite 3 mm	123	126
vite40	Vite 4 mm	126	127
dado30	Dado 3 mm	122	123
dado40	Dado 4 mm	126	127
rond50	Rondella 5 mm	123	126

Movimenti

Codice	Data	Carico	Scarico	CodFor	CodCli
vite40	01/01/2012	1200		124	
vite30	01/01/2012		800		825
vite30	02/01/2012	1000		954	
vite30	03/01/2012	2000		127	
rond50	03/01/2012		500		954

Fornitori

CodFor	Ditta	Indirizzo	Telefono
127	Vitoni spa	Via Ferri 2	0123,45678
122	Ferroni spa	Via Metalli 34	0234,5678
126	Nuova Metal	Via Industrie	0345,6789
123	Viti e Bulloni	Via di sopra 7	0567,9875

Clienti

CodCli	Ditta	Indirizzo	Telefono
925	Tendoni Max	Via di sotto 2	0113,44578
825	Arti Plus	Via di lato 45	0765,23456

La prima relazione, '**`Articoli`**', rappresenta l'anagrafica del magazzino di un grossista di ferramenta. Ogni articolo di magazzino viene codificato e descritto, inoltre vengono annotati i riferimenti ai codici di possibili fornitori. La seconda relazione, '**`Movimenti`**', elenca le operazioni di carico e di scarico degli articoli di magazzino, specificando solo il codice dell'articolo, la data, la quantità caricata o scaricata e il codice del fornitore o del cliente da cui è stato acquistato o a cui è stato venduto l'articolo. Infine seguono le relazioni che descrivono i codici dei fornitori e quelli dei clienti.

Si può intendere che una sola relazione non avrebbe potuto essere utilizzata utilmente per esprimere tutte queste informazioni.

È importante stabilire che, nel modello relazionale, il collegamento tra le tuple delle varie relazioni avviene attraverso dei valori e non

attraverso dei puntatori. Infatti, nella relazione '**Articoli**' l'attributo '**Fornitore1**' contiene il valore 123 e questo significa solo che i dati di quel fornitore sono rappresentati da quel valore. Nella relazione '**Fornitori**', la tupla il cui attributo '**CodFor**' contiene il valore 123 è quella che contiene i dati di quel particolare fornitore. Quindi, «123» non rappresenta un puntatore, ma solo una tupla che contiene quel valore nell'attributo «giusto». In questo senso si ribadisce l'indifferenza della posizione delle tuple all'interno delle relazioni.

74.1.2.2 Tipi di dati, domini e informazioni mancanti

Nelle relazioni, ogni attributo contiene una singola informazione elementare di un certo tipo, per il quale esiste un dominio determinato di valori possibili. Ogni attributo di ogni tupla deve contenere un valore ammissibile, nell'ambito del proprio dominio.

Spesso capitano situazioni in cui i valori di uno o più attributi di una tupla non sono disponibili per qualche motivo. In tal caso si pone il problema di assegnare a questi attributi un valore che definisca in modo non ambiguo questo stato di indeterminatezza. Questo valore viene definito come '**NULL**' ed è ammissibile per tutti i tipi di attributi possibili.

74.1.2.3 Vincoli di validità

I dati contenuti in una o più relazioni sono utili in quanto «sensati» in base al contesto a cui si riferiscono. Per esempio, considerando la relazione '**Movimenti**', vista precedentemente, questa deve contenere sempre un codice valido nell'attributo '**Codice**'. Se così non fosse, la registrazione data da quella tupla che dovesse avere un riferimen-

to a un codice di articolo non valido, non avrebbe alcun senso, perché mancherebbe proprio l'informazione più importante: l'articolo caricato o scaricato.

Il controllo sulla validità dei dati può avvenire a diversi livelli, a seconda della circostanza. Si possono distinguere vincoli che riguardano:

1. il dominio dell'attributo stesso -- quando si tratta di definire se l'attributo può assumere il valore '`NULL`' o meno e quando si stabilisce l'intervallo dei valori ammissibili;
2. gli altri attributi della stessa tupla -- quando dal valore contenuto in altri attributi della stessa tupla dipende l'intervallo dei valori ammissibili per l'attributo in questione;
3. gli attributi di altre tuple -- quando dal valore contenuto negli attributi delle altre tuple della stessa relazione dipende l'intervallo dei valori ammissibili per l'attributo in questione;
4. gli attributi di tuple di altre relazioni -- quando altre relazioni condizionano la validità di un attributo determinato.

I vincoli di tupla, ovvero quelli che riguardano i primi due punti dell'elenco appena indicato, sono i più semplici da esprimere perché non occorre conoscere altre informazioni esterne alla tupla stessa. Per esempio, un attributo che esprime un prezzo potrebbe essere definito in modo tale che non sia ammissibile un valore negativo; nello stesso modo, un attributo che esprime uno sconto su un prezzo potrebbe ammettere un valore positivo diverso da zero solo se il prezzo a cui si riferisce, contenuto nella stessa tupla, supera un valore determinato.

Il caso più importante di un vincolo interno alla relazione, che coinvolge più tuple, è quello che riguarda le ***chiavi***. In certe situazioni, un attributo, o un gruppo particolare di attributi di una relazione, deve essere unico per ogni tupla. Quindi, questo attributo, o questo gruppo di attributi, è valido solo quando non è già presente in un'altra tupla della stessa relazione.

Quando le informazioni sono distribuite fra più relazioni, i dati sono validi solo quando tutti i riferimenti sono validi. Volendo riprendere l'esempio della gestione di magazzino, visto precedentemente, una tupla della relazione '**`Movimenti`**' che dovesse contenere un codice di un fornitore o di un cliente inesistente, si troverebbe, in pratica, senza questa informazione.

74.1.2.4 Chiavi

Nella sezione precedente si è accennato alle ***chiavi***. Questo concetto merita un po' di attenzione. In precedenza è stato affermato che una relazione contiene una raccolta di tuple che contano per il loro contenuto e non per la loro posizione. In questo senso non è ammissibile una relazione contenente due tuple identiche. Una ***chiave*** di una relazione è un gruppo di attributi che permette di identificare univocamente le tuple in essa contenute; per questo, tali attributi devono contenere dati differenti per ogni tupla.

Stabilendo quali attributi devono costituire una chiave per una certa relazione, si comprende intuitivamente che questi attributi non possono mai contenere un valore indeterminato.

Nella definizione di relazioni collegate attraverso dei riferimenti, l'oggetto di questi riferimenti deve essere una chiave per la relazio-

ne di destinazione. Diversamente non si otterrebbe un riferimento univoco a una tupla particolare.

74.1.3 Gestione delle relazioni

«

Prima di affrontare l'utilizzo pratico di una base di dati relazionale, attraverso un linguaggio di manipolazione dei dati, è opportuno considerare a livello teorico alcuni tipi di operazioni che si possono eseguire con le relazioni.

Inizialmente è stato affermato che una relazione è un insieme di tuple... Dalla teoria degli insiemi derivano molte delle operazioni che riguardano le relazioni.

74.1.3.1 Unione, intersezione e differenza

«

Quando si maneggiano relazioni contenenti gli stessi attributi, hanno senso le operazioni fondamentali sugli insiemi: unione, intersezione e differenza. Il significato è evidente: l'unione genera una relazione composta da tutte le tuple distinte delle relazioni di origine; l'intersezione genera una relazione composta dalle tuple presenti simultaneamente in tutte le relazioni di origine; la differenza genera una relazione contenente le tuple che compaiono esclusivamente nella prima delle relazioni di origine.

L'esempio rappresentato dalle relazioni della figura 74.4 dovrebbe chiarire il senso di queste affermazioni.

Figura 74.4. Unione, intersezione e differenza tra relazioni.

Laureati

Codice	Nominativo	...
1245	Tizi Tizio	...
1745	Cai Caio	...
1655	Semproni Sempronio	...

Magazzinieri

Codice	Nominativo	...
1745	Cai Caio	...
1986	Pallino Pinco	...
1245	Tizi Tizio	...

Laureati UNITO Magazzinieri

Codice	Nominativo	...
1245	Tizi Tizio	...
1745	Cai Caio	...
1655	Semproni Sempronio	...
1986	Pallino Pinco	...

Laureati INTERSECATO Magazzinieri

Codice	Nominativo	...
1245	Tizi Tizio	...
1745	Cai Caio	...

Laureati MENO Magazzinieri

Codice	Nominativo	...
1655	Semproni Sempronio	...

74.1.3.2 Ridenominazione degli attributi

«

L'elaborazione dei dati contenuti in una relazione può avvenire previa modifica dei nomi di alcuni attributi. La modifica dei nomi genera di fatto una nuova relazione temporanea, per il tempo necessario a eseguire l'elaborazione conclusiva.

Le situazioni in cui la ridenominazione degli attributi può essere conveniente possono essere varie. Nel caso delle operazioni sugli insiemi visti nella sezione precedente, la ridenominazione può rendere compatibili relazioni i cui attributi, pur essendo compatibili, hanno nomi differenti.

74.1.3.3 Selezione, proiezione e congiunzione

«

La ***selezione*** e la ***proiezione*** sono operazioni che si eseguono su una sola relazione e generano una relazione che contiene una porzione dei dati di quella di origine. La selezione permette di estrarre alcune tuple dalla relazione, mentre la proiezione estrae parte degli attributi di tutte le tuple. Il primo caso, quello della selezione, non richiede considerazioni particolari, mentre la proiezione ha delle implicazioni importanti.

Attraverso la proiezione, utilizzando solo parte degli attributi, si genera una relazione in cui si potrebbero perdere delle tuple, a causa della possibilità che risultino dei doppioni. Per esempio, si consideri la relazione mostrata nella figura 74.5.

Figura 74.5. Relazione '**`Utenti (UID, Nominativo, Cognome, Nome, Ufficio)`**'.

Utenti				
UID	**Nominativo**	**Cognome**	**Nome**	**Ufficio**
0	root	Pallino	Pinco	CED
515	rmario	Rossi	Mario	Contabilità
501	bbianco	Bianchi	Bianco	Magazzino
502	rrosso	Rossi	Rosso	Contabilità

La figura mostra una relazione contenente le informazioni sugli utenti di un centro di elaborazione dati. Si può osservare che sia l'attributo '**`UID`**', sia l'attributo '**`Nominativo`**', possono essere da soli una chiave per la relazione. Se da questa relazione si vuole ottenere una proiezione contenente solo gli attributi '**`Cognome`**' e '**`Ufficio`**', non essendo questi due una chiave della relazione, si perdono delle tuple.

Figura 74.6. Proiezione degli attributi '**`Cognome`**' e '**`Ufficio`**' della relazione '**`Utenti`**'.

Cognome	**Ufficio**
Pallino	CED
Rossi	Contabilità
Bianchi	Magazzino

La figura 74.6 mostra la proiezione della relazione '**`Utenti`**', in cui

sono stati estratti solo gli attributi '**`Cognome`**' e '**`Ufficio`**'. In tal modo, le tuple che prima corrispondevano al numero '**`UID`**' 515 e 502 si sono ridotte a una sola, quella contenente il cognome «Rossi» e l'ufficio «Contabilità».

Da questo esempio si dovrebbe intendere che la proiezione ha senso, prevalentemente, quando gli attributi estratti costituiscono una chiave della relazione originaria

La ***congiunzione*** di relazioni, o *join*, è un'operazione in cui due o più relazioni vengono unite a formare una nuova relazione. Questo congiungimento implica la creazione di tuple formate dall'unione di tuple provenienti dalle relazioni di origine. Se per semplicità si pensa solo alla congiunzione di due relazioni: si va da una congiunzione minima in cui nessuna tupla della prima relazione risulta abbinata ad altre tuple della seconda, fino a un massimo in cui si ottengono tutti gli abbinamenti possibili delle tuple della prima relazione con quelle della seconda. Tra questi estremi si pone la situazione tipica, quella in cui ogni tupla della prima relazione viene collegata solo a una tupla corrispondente della seconda.

La ***congiunzione naturale*** si ottiene quando le relazioni oggetto di tale operazione vengono collegate in base ad attributi aventi lo stesso nome. Per osservare di cosa si tratta, vale la pena di riprendere l'esempio della gestione di magazzino già descritto in precedenza. Nella figura 74.7 ne viene mostrata nuovamente solo una parte.

Figura 74.7. Le relazioni '**`Articoli`**' e '**`Movimenti`**' dell'esempio sulla gestione del magazzino.

Articoli		
Codice	**Descrizione**	**...**
vite30	Vite 3 mm	...
vite40	Vite 4 mm	...
dado30	Dado 3 mm	...
dado40	Dado 4 mm	...
rond50	Rondella 5 mm	...

Movimenti				
Codice	**Data**	**Carico**	**Scarico**	**...**
vite40	01/01/2012	1200		...
vite30	01/01/2012		800	...
vite30	02/01/2012	1000		...
vite30	03/01/2012	2000		...
rond50	03/01/2012		500	...

La congiunzione naturale delle relazioni '**`Movimenti`**' e '**`Articoli`**', basata sulla coincidenza del contenuto dell'attributo '**`Codice`**', genera una relazione in cui appaiono tutti gli attributi delle due relazioni di origine, con l'eccezione dell'attributo '**`Codice`**' che appare una volta sola (figura 74.8).

Tabella 74.8. Il join naturale tra le relazioni '**Movimenti**' e '**Articoli**'.

Codice	**Data**	**Carico**	**Scarico**	**...**	**Descrizione**	**...**
vite40	01/01/2012	1200		...	Vite 4 mm	...
vite30	01/01/2012		800	...	Vite 3 mm	...
vite30	02/01/2012	1000		...	Vite 3 mm	...
vite30	03/01/2012	2000		...	Vite 3 mm	...
rond50	03/01/2012		500	...	Rondella 5 mm	...

Nel caso migliore, ogni tupla di una relazione trova una tupla corrispondente dell'altra; nel caso peggiore, nessuna tupla ha una corrispondente nell'altra relazione. L'esempio mostra che tutte le tuple della relazione '**Movimenti**' hanno trovato una corrispondenza nella relazione '**Articoli**', mentre solo alcune tuple della relazione '**Articoli**' hanno una corrispondenza dall'altra parte. Queste tuple, quelle che non hanno una corrispondenza, sono dette «penzolanti», o *dangling*, e di fatto vengono perdute dopo la congiunzione.

Quando una congiunzione genera corrispondenze per tutte le tuple delle relazioni coinvolte, si parla di congiunzione completa. La dimensione (espressa in quantità di tuple) della relazione risultante in presenza di una congiunzione completa è pari alla dimensione massima delle varie relazioni.

Quando si vuole eseguire la congiunzione di relazioni che non hanno attributi in comune si ottiene il collegamento di ogni tupla di una relazione con ogni tupla delle altre. Si può osservare l'esempio della

figura 74.9 che riprende il solito problema del magazzino, con delle semplificazioni opportune.

Figura 74.9. Le relazioni '**`Articoli`**' e '**`Fornitori`**' dell'esempio sulla gestione del magazzino.

Articoli			
Codice	**Descrizione**	**Fornitore1**	**...**
vite30	Vite 3 mm	127	...
vite40	Vite 4 mm	127	...
dado30	Dado 3 mm	122	...

Fornitori		
CodFor	**Ditta**	**...**
127	Vitoni spa	...
122	Ferroni spa	...

Nessuno degli attributi delle due relazioni coincide, quindi si ottiene un «prodotto» tra le due relazioni, in pratica, una relazione che contiene il prodotto delle tuple contenute nelle relazioni originali (figura 74.10).

Figura 74.10. Il prodotto tra le relazioni '**Articoli**' e '**Fornitori**'.

Codice	Descrizione	Fornitore1	...	CodFor	Ditta	...
vite30	Vite 3 mm	127	...	127	Vitoni spa	...
vite40	Vite 4 mm	127	...	127	Vitoni spa	...
dado30	Dado 3 mm	122	...	127	Vitoni spa	...
vite30	Vite 3 mm	127	...	122	Ferroni spa	...
vite40	Vite 4 mm	127	...	122	Ferroni spa	...
dado30	Dado 3 mm	122	...	122	Ferroni spa	...

Quando si esegue un'operazione del genere, è normale che molte delle tuple risultanti siano prive di significato per gli scopi che ci si prefigge. Di conseguenza, quasi sempre, si applica poi una selezione attraverso delle condizioni. Nel caso dell'esempio, sarebbe ragionevole porre come condizione di selezione l'uguaglianza tra i valori dell'attributo '**Fornitore1**' e '**CodFor**'.

Figura 74.11. La selezione delle tuple che rispettano la condizione di uguaglianza tra gli attributi '**Fornitore1**' e '**CodFor**'.

Codice	Descrizione	Fornitore1	...	CodFor	Ditta	...
vite30	Vite 3 mm	127	...	127	Vitoni spa	...
vite40	Vite 4 mm	127	...	127	Vitoni spa	...
dado30	Dado 3 mm	122	...	122	Ferroni spa	...

Generalmente, nella pratica, non esiste la possibilità di definire una congiunzione basato sull'uguaglianza dei nomi degli attributi. Di solito si esegue una congiunzione che genera un prodotto tra le relazioni, quindi si applicano delle condizioni di selezione come nell'esempio mostrato. Quando la selezione in questione è del tipo visto nell'esempio, cioè basata sull'uguaglianza del contenuto di attributi delle diverse relazioni (anche se il nome di questi attributi è differente), si parla di ***equi-giunzione*** (*equi-join*).

74.1.3.4 Gestione dei valori nulli

Si è accennato in precedenza alla possibilità che gli attributi di una relazione possano contenere anche il valore indeterminato, o '**`NULL`**'. Con questo valore indeterminato non si possono fare comparazioni con valori determinati e di solito nemmeno con altri valori indeterminati. Per esempio, non si può dire che '**`NULL`**' sia maggiore o minore di qualcosa; una comparazione di questo tipo genera solo un risultato indeterminato. '**`NULL`**' è solo uguale a se stesso ed è diverso da ogni altro valore, compreso un altro valore '**`NULL`**'.

Per verificare la presenza o l'assenza di un valore indeterminato si utilizzano generalmente operatori specifici, come in SQL:

- '**`IS NULL`**' -- che si avvera quando il valore controllato è indeterminato;
- '**`IS NOT NULL`**' -- che si avvera quando il valore controllato è determinato, quindi diverso da indeterminato.

Nel momento in cui si eseguono delle espressioni logiche, utilizzando i soliti operatori AND, OR e NOT, si pone il problema di stabilire cosa accade quando si presentano valori indeterminati. La soluzione

è intuitiva: quando non si può fare a meno di conoscere il valore che si presenta come indeterminato, il risultato è indeterminato. Questo concetto deriva dalla cosiddetta logica *fuzzy*.

Figura 74.12. Tabella della verità degli operatori AND e OR quando sono coinvolti valori indefiniti.

```
? AND ? = ?      ? OR ? = ?
? AND F = F      ? OR F = ?
? AND T = ?      ? OR T = T
F AND ? = F      F OR ? = ?
F AND F = F      F OR F = F
F AND T = F      F OR T = T
T AND ? = ?      T OR ? = T
T AND F = F      T OR F = T
T AND T = T      T OR T = T
```

74.1.3.5 Relazioni derivate e viste

«

Precedentemente si è accennato al fatto che la rappresentazione finale dei dati può essere diversa da quella logica. Nel modello relazionale è possibile ottenere delle relazioni derivate da altre, attraverso una funzione determinata che stabilisce il modo con cui ottenere queste derivazioni. Si distingue fondamentalmente tra:

- relazioni derivate virtuali, o ***viste***, che non generano nuove relazioni memorizzate nella base di dati, il cui contenuto viene generato al volo al momento della necessità;
- relazioni derivate materializzate, che generano una nuova relazione nella base di dati.

Il primo dei due casi è semplice da gestire, perché i dati sono sempre allineati correttamente, ma è pesante dal punto di vista elaborativo;

il secondo ha invece i pregi e i difetti opposti. Con il termine «vista» si intende fare riferimento alle relazioni derivate virtuali.

74.2 Questioni organizzative generali dei DBMS comuni

Dopo la teoria astratta, l'organizzazione di un DBMS richiede una realizzazione pratica, che implica delle scelte. In questa sezione si descrivono alcune questioni «pratiche» che è bene conoscere inizialmente, prima di affrontare lo studio di un DBMS specifico.

74.2.1 Configurazione e basi di dati amministrative

Un DBMS, per funzionare, deve poter annotare l'esistenza e i contenuti delle proprie basi di dati, così come l'esistenza e i privilegi dei propri utenti. Per conservare queste informazioni, può gestire dei file di configurazione, oppure, più spesso, impiegare una o più basi di dati amministrative, la cui gestione avviene in modo quasi trasparente.

La presenza di queste basi di dati amministrative implica il fatto che non possano esserne create delle altre con gli stessi nomi, ma ci sono naturalmente anche altre implicazioni più importanti.

Quando si installa il software di gestione del DBMS per la prima volta, è necessario provvedere a costruire le basi di dati amministrative, oltre che, eventualmente, a sistemare altri file di configurazione. Per questo, di solito il software del DBMS include un programma che predispone tali basi di dati speciali con una configurazione iniziale predefinita.

Quando si aggiorna il software di gestione del DBMS, si ha generalmente la necessità di conservare i dati preesistenti. Questo signi-

fica preservare le basi di dati che sono state create e le informazioni sulle utenze, con i privilegi rispettivi. Il problema sta nel fatto che il software aggiornato potrebbe avere un'organizzazione differente nel modo di gestire i file che contengono le informazioni sulle basi di dati, pertanto è poi necessario provvedere a una conversione, con l'ausilio di strumenti realizzati appositamente per quel DBMS. Naturalmente, per evitare circoli viziosi, un software aggiornato dovrebbe essere in grado di accedere a basi di dati di qualche versione precedente.

Generalmente, prima di un aggiornamento del software di gestione del DBMS, si consiglia di eseguire una copia dei dati in una forma indipendente dalla versione, che può essere acquisita successivamente, dopo l'aggiornamento; tuttavia, rimane il problema delle basi di dati amministrative: se dovesse fallire la procedura di aggiornamento automatica, si renderebbe necessaria, nuovamente, la creazione delle basi di dati (vuote) e delle utenze.

74.2.2 Copie di sicurezza delle basi di dati

«

Generalmente, le copie di sicurezza del contenuto di una basi di dati, si fanno in modo di generare del codice SQL, contenente le istruzioni per la creazione delle relazioni e per l'inserimento delle tuple. Questo viene ottenuto tramite programmi o script del software del DBMS e il codice che si ottiene è specifico di quel tipo di DBMS, perciò non è universale.

Ci possono essere dei DBMS che consentono l'acquisizione di dati molto complessi in un solo attributo (dove per «attributo» si intende una cella di una riga di una tabella), ma poi, questi dati non possono essere rappresentati in modo testuale in un codice SQL. In tali casi,

l'archiviazione in forma di codice SQL si deve limitare ai dati consueti, ignorando il resto; pertanto si devono usare formati differenti per l'archiviazione completa dei dati.

74.2.3 Comunicazione con il DBMS e accesso alle basi di dati

I programmi accedono alle basi di dati attraverso un protocollo di comunicazione con il DBMS. Il protocollo in questione dipende dal DBMS, ma generalmente consente di trasportare delle istruzioni SQL.

Nei sistemi Unix, la comunicazione con il DBMS avviene tipicamente attraverso socket di dominio Unix, per le comunicazioni locali, e socket di dominio Internet per quelle remote. Pertanto, il DBMS deve disporre di un demone in attesa di dati da un file di tipo socket e in ascolto di una porta TCP o UDP (di solito si tratta del protocollo TCP).

Naturalmente, per consentire l'accesso alle basi di dati, il DBMS deve avere un modo per «riconoscere» chi vuole accedere.

Un DBMS che consente esclusivamente collegamenti di tipo locale, avrebbe la possibilità di individuare gli accessi in base al numero UID associato al processo elaborativo del programma che tenta il contatto. In questo caso particolare, il riconoscimento delle utenze può essere demandata al sistema operativo.

74.2.4 Utenze e privilegi

Le utenze di un DBMS servono a distinguere le competenze al suo interno. Generalmente si distingue la presenza di un amministratore con poteri illimitati nell'ambito della gestione complessiva del DBMS e di conseguenza di ogni base di dati. Il nome, dal punto del DBMS, di questo amministratore, non è standardizzato. A titolo di esempio, nel caso di PostgreSQL si tratta normalmente dell'utente '**`postgres`**', mentre con MySQL si usa il nome '**`root`**'.

Generalmente, il DBMS riconosce gli utenti attraverso una parola d'ordine, che deve essere fornita all'inizio di ogni collegamento; tuttavia, dovrebbe esistere anche la possibilità di definire utenze senza parola d'ordine (sulla fiducia), oppure dovrebbe essere possibile definire dei contesti per cui l'accesso non debba richiedere questa formalità.

Il problema di evitare l'obbligo di inserire la parola d'ordine si sente in particolare proprio per l'accesso in qualità di amministratore, quando si vogliono realizzare degli script per svolgere certe funzioni amministrative. Le forme di raggiro dipendono dalle caratteristiche del DBMS.

Quando il DBMS è in grado di riconoscere un accesso locale, in quanto proveniente da un utente che ha lo stesso nome usato nell'ambito del sistema operativo, potrebbe accettarlo senza richiesta di una parola d'ordine, perché in sostanza l'autenticazione è già avvenuta. Questo meccanismo viene usato in particolare con PostgreSQL, dove l'utente '**`postgres`**' viene aggiunto anche nel file '`/etc/passwd`'; in tal modo, ammesso che la configurazione di PostgreSQL lo consenta, l'utente '**`root`**' del sistema operativo

può impersonare facilmente l'utente '**postgres**', attraverso il comando '**su**', quindi può accedere localmente al DBMS venendo riconosciuto implicitamente.

Quando non c'è la possibilità di sfruttare il sistema operativo per il riconoscimento dell'utente in modo implicito, si può arrivare ad annotare la parola d'ordine (in chiaro) in un file che solo l'utente '**root**' del sistema operativo può leggere, così che uno script con i privilegi necessari possa leggere questa parola d'ordine, usarla per collegarsi al DBMS e svolgere il suo compito.

Generalmente, le utenze vengono considerate nel DBMS soltanto come nominativi puri e semplici, senza distinguerne la provenienza. Il DBMS potrebbe disporre di una configurazione ulteriore in cui si specifica il metodo di riconoscimento richiesto, in base alla provenienza degli accessi (PostgreSQL), oppure potrebbe arrivare a considerare le utenze come l'unione del nome al nodo di origine, come se si trattasse di utenti distinti. Questo secondo caso riguarda in particolare MySQL e vale la pena di considerarlo con attenzione, perché si possono creare degli equivoci; infatti, se un'utenza è abbinata all'origine '**localhost**' e tale utente accede sì dall'elaboratore locale (come indica convenzionalmente il nome '**localhost**'), ma il file '`/etc/hosts`' non abbina correttamente l'indirizzo locale a tale nome, l'accesso fallisce; inoltre, se l'utenza fosse abbinata all'origine '**127.0.0.1**' e l'utente cercasse di accedere localmente, potrebbe succedergli di non essere riconosciuto se il sistema operativo traduce poi l'indirizzo nel nome.

Per quanto riguarda la gestione delle utenze, c'è da considerare che esistono anche dei DBMS semplificati che, concedendo esclusivamente accessi locali, invece di gestire le utenze in proprio si affi-

dano alla gestione del sistema operativo. In questi casi, i permessi di accesso alle basi di dati vengono regolati tramite la gestione dei permessi corrispondenti ai file che rappresentano le basi di dati stesse.

74.2.5 ODBC

«

ODBC, ovvero *Open database connectivity* è un metodo standardizzato per l'accesso ai DBMS. In pratica, si inserisce un servizio intermedio, tra i DBMS e le applicazioni che devono accedere ai dati: le applicazioni comunicano con il servizio ODBC; il servizio ODBC comunica con i DBMS sottostanti, preoccupandosi di adattarsi alle loro particolarità. In questo modo, invece di scrivere applicazioni che comunicano solo con un certo DBMS, le applicazioni fatte per ODBC, possono utilizzare qualsiasi DBMS che il servizio ODBC è in grado di gestire.

74.3 Introduzione a SQL

«

SQL è l'acronimo di *Structured query language* e identifica un linguaggio di interrogazione (gestione) per basi di dati relazionali. Le sue origini risalgono alla fine degli anni 1970 e questo giustifica la sua sintassi prolissa e verbale tipica dei linguaggi dell'epoca, come il COBOL.

Allo stato attuale, data la sua evoluzione e standardizzazione, l'SQL rappresenta un riferimento fondamentale per la gestione di una base di dati relazionale.

A parte il significato originale dell'acronimo, SQL è un linguaggio completo per la gestione di una base di dati relazionale, includendo

le funzionalità di un DDL (*Data definition language*), di un DML (*Data manipulation language*) e di un DCL (*Data control language*).

Data l'età e la conseguente evoluzione di questo linguaggio, si sono definiti nel tempo diversi livelli di standard. I più importanti sono SQL89, SQL92 e SQL99, noti anche come SQL1, SQL2 e SQL3 rispettivamente.

L'aderenza dei vari sistemi DBMS allo standard SQL92 non è mai completa e perfetta, per questo sono stati definiti dei sottolivelli di questo standard per definire il grado di compatibilità di un DBMS. Si tratta di: *entry SQL*, *intermediate SQL* e *full SQL*. Si può intendere che il primo sia il livello di compatibilità minima e l'ultimo rappresenti la compatibilità totale. Lo standard di fatto è rappresentato prevalentemente dal primo livello, che coincide fondamentalmente con lo standard precedente, SQL89. Da questo si comprende che lo stato di assimilazione di SQL99 è ancora più arretrato.

74.3.1 Concetti fondamentali

Convenzionalmente, le istruzioni di questo linguaggio sono scritte con tutte le lettere maiuscole. Si tratta solo di una tradizione di quell'epoca. SQL non distingue tra lettere minuscole e maiuscole nelle parole chiave delle istruzioni e nemmeno nei nomi di relazioni, attributi e altri oggetti. Solo quando si tratta di definire il contenuto di una variabile, allora le differenze contano.

In questo capitolo e nel resto del documento, quando si fa riferimento a istruzioni SQL, queste vengono indicate utilizzando solo lettere maiuscole, come richiede la tradizione.

I nomi degli oggetti (relazioni e altro) possono essere composti utilizzando lettere, numeri e il trattino basso; il primo carattere deve

essere una lettera oppure il trattino basso.

Le istruzioni SQL possono essere distribuite su più righe, senza una regola precisa. Si distingue la fine di un'istruzione dall'inizio di un'altra attraverso la presenza di almeno una riga vuota. Generalmente, i sistemi SQL richiedono l'uso di un simbolo di terminazione delle righe, che di norma è costituito dal punto e virgola.

L'SQL standard prevede la possibilità di inserire commenti; per questo si può usare un trattino doppio ('--') seguito dal commento desiderato, fino alla fine della riga. Tuttavia, si osservi che per ottenere la massima compatibilità con i DBMS esistenti, conviene lasciare almeno uno spazio dopo il trattino doppio, prima di inserire il commento vero e proprio.

74.3.2 Terminologia

«

Il linguaggio SQL utilizza una propria terminologia per distinguere gli «oggetti» che manipola. Per la precisione, si utilizzano normalmente i termini «tabella», «riga» e «colonna», al posto di «relazione», «tupla» e «attributo».

Esistono delle buone ragioni per utilizzare una terminologia differente nel linguaggio SQL, soprattutto in considerazione del fatto che in questo caso sono ammissibili situazioni che nella teoria generale delle basi di dati relazionali non lo sarebbero (per esempio la possibilità di avere tuple doppie). Tuttavia, si osservi che in questo documento si cerca di mantenere una certa uniformità nei termini, seguendo la tradizione della teoria delle basi di dati, anche a costo di rischiare una contraddizione con questa.

Specchietto 74.13. Associazione tra i termini relativi alla gestione delle basi di dati. Ogni riga dello specchietto, rappresenta un contesto differente, mentre le colonne individuano la traduzione dei termini in base al contesto.

classi	istanze	attributi	tipi di dati contenibili negli attributi
relazioni	tuple	attributi	domini
tabelle	righe	colonne	tipi di dati contenibili nelle colonne

74.4 DDL

DDL, ovvero *Data definition language*, è il linguaggio usato per definire la struttura dei dati (in una base di dati). In questa sezione viene trattato il linguaggio SQL per ciò che riguarda specificatamente i dati, la loro creazione e la loro distruzione.

74.4.1 Tipi di dati

I tipi di dati gestibili con il linguaggio SQL sono molti. Fondamentalmente si possono distinguere tipi contenenti: valori numerici, stringhe e informazioni data-orario. Nelle sezioni seguenti vengono descritti solo alcuni dei tipi definiti dallo standard.

74.4.1.1 Stringhe di caratteri

Si distinguono due tipi di stringhe di caratteri in SQL: quelle a dimensione fissa, completate a destra dal carattere spazio, e quelle a dimensione variabile.

Sintassi	Descrizione

`CHARACTER`
 	
`CHARACTER(`***dimensione***`)`
 	
`CHAR`
 	
`CHAR(`***dimensione***`)`
	Queste sono le varie sintassi alternative che possono essere utilizzate per definire una stringa di dimensione fissa. Se non viene indicata la dimensione tra parentesi, si intende una stringa di un solo carattere.

`CHARACTER VARYING(`***dimensione***`)`
 	
`CHAR VARYING(`***dimensione***`)`
 	
`VARCHAR(`***dimensione***`)`
	Una stringa di dimensione variabile può essere definita attraverso uno dei tre modi appena elencati. È necessario specificare la dimensione massima che questa stringa può avere. Il minimo è rappresentato dalla stringa nulla.

Le costanti stringa si esprimono delimitandole attraverso apici singoli, oppure apici doppi, come nell'esempio seguente:

```
'Questa è una stringa letterale per SQL'
"Anche questa è una stringa letterale per SQL"
```

Non tutti i sistemi SQL accettano entrambi i tipi di delimitatori di stringa. In caso di dubbio è bene limitarsi all'uso degli apici singoli; eventualmente, per inserire un apice singolo in una stringa delimitata con apici singoli, dovrebbe essere sufficiente il suo raddoppio. In pratica, per scrivere una stringa del tipo «l'albero», dovrebbe essere possibile scrivere:

```
'l''albero'
```

74.4.1.2 Valori numerici

I tipi numerici si distinguono in ***esatti*** e ***approssimati***, intendendo con la prima definizione quelli di cui si conosce il numero massimo di cifre numeriche intere e decimali, mentre con la seconda si fa riferimento ai tipi a virgola mobile. In ogni caso, le dimensioni massime o la precisione massima che possono avere tali valori dipende dal sistema in cui vengono utilizzati.

Sintassi	Descrizione

`NUMERIC`
 	
`NUMERIC(`***precisione*** [`,` ***scala***] `)`
	Il tipo '**`NUMERIC`**' permette di definire un valore numerico composto da un massimo di tante cifre numeriche quante indicate dalla precisione, cioè il primo argomento tra parentesi. Se viene specificata anche la scala, si intende riservare quella parte di cifre per quanto appare dopo la virgola. Per esempio, con '**`NUMERIC(5,2)`**' si possono rappresentare valori da +999,99 a −999,99.
	Se non viene specificata la scala, si intende che si tratti solo di valori interi; se non viene specificata nemmeno la precisione, viene usata la definizione predefinita per questo tipo di dati, che dipende dalle caratteristiche del DBMS.

Sintassi	Descrizione

`DECIMAL`
`DECIMAL(`***precisione*** [`,` ***scala***] `)`
`DEC`
`DEC(`***precisione*** [`,` ***scala***] `)`
	Il tipo '**`DECIMAL`**' è simile al tipo '**`NUMERIC`**', con la differenza che le caratteristiche della precisione e della scala rappresentano le esigenze minime, mentre il sistema può fornire una rappresentazione con precisione o scala maggiore.

`INTEGER`
`INT`
`SMALLINT`
	I tipi '**`INTEGER`**' e '**`SMALLINT`**' rappresentano tipi interi la cui dimensione dipende generalmente dalle caratteristiche del sistema operativo e dall'hardware utilizzato. L'unico riferimento sicuro è che il tipo '**`SMALLINT`**' permette di rappresentare interi con una dimensione inferiore o uguale al tipo '**`INTEGER`**'.

`FLOAT` `FLOAT(`***precisione***`)`	Il tipo '**`FLOAT`**' definisce un tipo numerico approssimato (a virgola mobile) con una precisione binaria pari o superiore di quella indicata tra parentesi (se non viene indicata, dipende dal sistema).
`REAL` `DOUBLE PRECISION`	Il tipo '**`REAL`**' e il tipo '**`DOUBLE PRECISION`**' sono due tipi a virgola mobile con una precisione prestabilita. Questa precisione dipende dal sistema, ma in generale, il secondo dei due tipi deve essere più preciso dell'altro.

I valori numerici costanti vengono espressi attraverso la semplice indicazione del numero senza delimitatori. La virgola di separazione della parte intera da quella decimale si esprime normalmente attraverso il punto (‘.’), a meno che sia prevista una forma di adattamento alle caratteristiche di configurazione locale.

74.4.1.3 Valori Data-orario

I valori data-orario sono di tre tipi e servono rispettivamente a memorizzare un giorno particolare, un orario normale e un’informazione data-ora completa.

Sintassi	Descrizione
`DATE`	Il tipo ‘**`DATE`**’ permette di rappresentare delle date composte dall’informazione anno-mese-giorno.

`TIME`
`TIME(`***precisione***`)`
`TIME WITH TIME ZONE`
`TIME(`***precisione***`) WITH TIME ZONE`
	Il tipo ‘**`TIME`**’ permette di rappresentare un orario particolare, composto da ore-minuti-secondi ed eventualmente frazioni di secondo. Se viene specificata la precisione, si intende definire un numero di cifre per la parte frazionaria dei secondi, altrimenti si intende che non debbano essere memorizzate le frazioni di secondo.

Sintassi	Descrizione

`TIMESTAMP`
`TIMESTAMP(`***precisione***`)`
`TIMESTAMP WITH TIME ZONE`
`TIMESTAMP(`***precisione***`) WITH TIME ZONE`
	Il tipo '**`TIMESTAMP`**' è un'informazione oraria più completa del tipo '**`TIME`**' in quanto prevede tutte le informazioni, dall'anno ai secondi, oltre alle eventuali frazioni di secondo.

Se viene specificata la precisione, si intende definire un numero di cifre per la parte frazionaria dei secondi, altrimenti si intende che non debbano essere memorizzate le frazioni di secondo.

L'aggiunta dell'opzione '**`WITH TIME ZONE`**' serve a specificare un tipo orario differente, che assieme all'informazione oraria aggiunge lo scostamento, espresso in ore e minuti, dell'ora locale dal tempo universale (UTC). Per esempio, 22:05:10+1:00 rappresenta le 22.05 e 10 secondi dell'ora locale italiana (durante l'inverno), mentre il tempo universale corrispondente sarebbe invece 21:05:10+0:00.

Le costanti che rappresentano informazioni data-orario sono espresse come le stringhe, delimitate tra apici. Il sistema DBMS potrebbe ammettere più forme differenti per l'inserimento di queste, ma i modi più comuni dovrebbero essere quelli espressi dagli esempi seguenti.

```
'2012-12-31'
'12/31/2012'
'31.12.2012'
```

Questi tre esempi rappresentano la stessa data: il 31 dicembre 1999. Per una questione di uniformità, dovrebbe essere preferibile il primo

di questi formati, corrispondente allo stile ISO 8601. Anche gli orari che si vedono sotto, sono aderenti allo stile ISO 8601; in particolare per il fatto che il fuso orario viene indicato attraverso lo scostamento dal tempo universale, invece che attraverso una parola chiave che definisca il fuso dell'ora locale.

```
'12:30:50+1.00'
'12:30:50.10'
'12:30:50'
'12:30'
```

Il primo di questa serie di esempi rappresenta un orario composto da ore, minuti e secondi, oltre all'indicazione dello scostamento dal tempo universale (per ottenere il tempo universale deve essere sottratta un'ora). Il secondo esempio mostra un orario composto da ore, minuti, secondi e centesimi di secondo. Il terzo e il quarto sono rappresentazioni normali, in particolare nell'ultimo è stata omessa l'indicazione dei secondi.

```
'2012-12-31 12:30:50+1.00'
'2012-12-31 12:30:50.10'
'2012-12-31 12:30:50'
'2012-12-31 12:30'
```

Gli esempi mostrano la rappresentazione di informazioni data-orario complete per il tipo '**`TIMESTAMP`**'. La data è separata dall'ora da uno spazio.

74.4.1.4 Intervalli di tempo

Quanto mostrato nella sezione precedente rappresenta un valore che indica un momento preciso nel tempo: una data o un orario, o entrambe le cose. Per rappresentare una durata, si parla di intervalli. Per l'SQL si possono gestire gli intervalli a due livelli di precisione:

anni e mesi; oppure giorni, ore, minuti, secondi ed eventualmente anche le frazioni di secondo. L'intervallo si indica con la parola chiave '**INTERVAL**', seguita eventualmente dalla precisione con cui questo deve essere rappresentato:

```
INTERVAL [unità_di_misura_data_orario [TO unità_di_misura_data_orario]]
```

In pratica, si può indicare che si tratta di un intervallo, senza specificare altro, oppure si possono definire una o due unità di misura che limitano la precisione di questo (pur restando nei limiti a cui si è già accennato). Tanto per fare un esempio concreto, volendo definire un intervallo che possa esprimere solo ore e minuti, si potrebbe dichiarare con: '**INTERVAL HOUR TO MINUTE**'. La tabella 74.22 elenca le parole chiave che rappresentano queste unità di misura.

Tabella 74.22. Elenco delle parole chiave che esprimono unità di misura data-orario.

Parola chiave	Significato
YEAR	Anni
MONTH	Mesi
DAY	Giorni
HOUR	Ore
MINUTE	Minuti
SECOND	Secondi

Si osservino i due esempi seguenti:

```
INTERVAL '12 HOUR 30 MINUTE 50 SECOND'
INTERVAL '12:30:50'
```

Queste due forme rappresentano entrambe la stessa cosa: una durata di 12 ore, 30 minuti e 50 secondi. In generale, dovrebbe essere preferibile la seconda delle due forme di rappresentazione.

```
INTERVAL '10 DAY 12 HOUR 30 MINUTE 50 SECOND'
INTERVAL '10 DAY 12:30:50'
```

Come prima, i due esempi che si vedono sopra sono equivalenti. Intuitivamente, si può osservare che non ci può essere un altro modo di esprimere una durata in giorni, senza specificarlo attraverso la parola chiave '**DAY**'.

Per completare la serie di esempi, si aggiungono anche i casi in cui si rappresentano esplicitamente quantità molto grandi, che di conseguenza sono approssimate al mese (come richiede lo standard SQL92):

```
INTERVAL '10 YEAR 11 MONTH'
INTERVAL '10 YEAR'
```

Gli intervalli di tempo possono servire per indicare un tempo trascorso rispetto al momento attuale. Per specificare espressamente questo fatto, si indica l'intervallo come un valore negativo, aggiungendo all'inizio un trattino (il segno meno).

```
INTERVAL '- 10 YEAR 11 MONTH'
```

L'esempio che si vede sopra, esprime precisamente 10 anni e 11 mesi fa.

74.4.2 Operatori, funzioni ed espressioni

«

SQL, pur non essendo un linguaggio di programmazione completo, mette a disposizione una serie di operatori e di funzioni utili per la realizzazione di espressioni di vario tipo.

74.4.2.1 Operatori aritmetici

«

Gli operatori che intervengono su valori numerici sono elencati nella tabella 74.27.

Tabella 74.27. Elenco degli operatori aritmetici.

Operatore e operandi	Descrizione
−op	Inverte il segno dell'operando.
op1 + *op2*	Somma i due operandi.
op1 − *op2*	Sottrae dal primo il secondo operando.
op1 * *op2*	Moltiplica i due operandi.
op1 / *op2*	Divide il primo operando per il secondo.
op1 % *op2*	Modulo: il resto della divisione tra il primo e il secondo operando.

Nelle espressioni, tutti i tipi numerici esatti e approssimati possono essere usati senza limitazioni. Dove necessario, il sistema provvede a eseguire le conversioni di tipo.

74.4.2.2 Operazioni con i valori data-orario e con intervalli di tempo

Le operazioni che si possono compiere utilizzando valori data-orario e valori che esprimono intervalli di tempo, hanno significato solo in alcune circostanze. La tabella 74.28 elenca le operazioni possibili e il tipo di risultato che si ottiene in base al tipo di operatori utilizzato.

Tabella 74.28. Operatori e operandi validi quando si utilizzano valori data-orario e valori che esprimono intervalli di tempo.

Operatore e operandi	Risultato
data_orario − ***data_orario***	Intervallo
data_orario + ***intervallo***	Data-orario
data_orario − ***intervallo***	Data-orario
intervallo + ***data_orario***	Data-orario
intervallo + ***intervallo***	Intervallo
intervallo − ***intervallo***	Intervallo
intervallo * ***numerico***	Intervallo
intervallo / ***numerico***	Intervallo
numerico * ***intervallo***	Intervallo

74.4.2.3 Operatori di confronto e operatori logici

«

Gli operatori di confronto determinano la relazione tra due operandi. Il risultato dell'espressione composta da due operandi posti a confronto è di tipo booleano: *Vero* o *Falso*. Gli operatori di confronto sono elencati nella tabella 74.29.

Tabella 74.29. Elenco degli operatori di confronto.

Operatore e operandi	Descrizione
op1 = ***op2***	*Vero* se gli operandi si equivalgono.
op1 <> ***op2***	*Vero* se gli operandi sono differenti.
op1 < ***op2***	*Vero* se il primo operando è minore del secondo.
op1 > ***op2***	*Vero* se il primo operando è maggiore del secondo.
op1 <= ***op2***	*Vero* se il primo operando è minore o uguale al secondo.
op1 >= ***op2***	*Vero* se il primo operando è maggiore o uguale al secondo.

Quando si vogliono combinare assieme diverse espressioni logiche si utilizzano gli operatori logici. Come in tutti i linguaggi di programmazione, si possono usare le parentesi tonde per raggruppare le espressioni logiche in modo da chiarire l'ordine di risoluzione. Gli operatori logici sono elencati nella tabella 74.30.

Tabella 74.30. Elenco degli operatori logici.

Operatore e operandi	Descrizione
`NOT` ***op***	Inverte il risultato logico dell'operando.
op1 `AND` ***op2***	*Vero* se entrambi gli operandi restituiscono il valore *Vero*.
op1 `OR` ***op2***	*Vero* se almeno uno degli operandi restituisce il valore *Vero*.

Il meccanismo di confronto tra due operandi numerici è evidente, mentre può essere meno evidente con le stringhe di caratteri. Per la precisione, il confronto tra due stringhe avviene senza tenere conto degli spazi finali, per cui, le stringhe '**`'ciao'`**' e '**`'ciao '`**' dovrebbero risultare uguali attraverso il confronto di uguaglianza con l'operatore '**`=`**'.

Con le stringhe, tuttavia, si possono eseguire dei confronti basati su modelli, attraverso gli operatori '**`IS LIKE`**' e '**`IS NOT LIKE`**'. Il modello può contenere dei metacaratteri rappresentati dal trattino basso ('_'), che rappresenta un carattere qualsiasi, e dal simbolo di percentuale ('`%`'), che rappresenta una sequenza qualsiasi di caratteri. La tabella 74.31 riassume quanto affermato.

Tabella 74.31. Espressioni sulle stringhe di caratteri.

Espressioni e modelli	Descrizione
stringa `IS LIKE` ***modello***	Restituisce *Vero* se il modello corrisponde alla stringa.
stringa `IS NOT LIKE` ***modello***	Restituisce *Vero* se il modello non corrisponde alla stringa.
_	Rappresenta un carattere qualsiasi.

Espressioni e modelli	Descrizione
`%`	Rappresenta una sequenza indeterminata di caratteri.

La presenza di valori indeterminati richiede la disponibilità di operatori di confronto in grado di determinarne l'esistenza. La tabella 74.32 riassume gli operatori ammissibili in questi casi.

Tabella 74.32. Espressioni di verifica dei valori indeterminati.

Operatori	Descrizione
espressione `IS NULL`	Restituisce *Vero* se l'espressione genera un risultato indeterminato.
espressione `IS NOT NULL`	Restituisce *Vero* se l'espressione non genera un risultato indeterminato.

Infine, occorre considerare una categoria particolare di espressioni che permettono di verificare l'appartenenza di un valore a un intervallo o a un elenco di valori. La tabella 74.33 riassume gli operatori utilizzabili.

Tabella 74.33. Espressioni per la verifica dell'appartenenza di un valore a un intervallo o a un elenco.

Operatori e operandi	Descrizione
op1 `IN (`***elenco***`)`	*Vero* se il primo operando è contenuto nell'elenco.
op1 `NOT IN (`***elenco***`)`	*Vero* se il primo operando non è contenuto nell'elenco.
op1 `BETWEEN` ***op2*** `AND` ***op3***	*Vero* se il primo operando è compreso tra il secondo e il terzo.
op1 `NOT BETWEEN` ***op2*** `AND` ***op3***	*Vero* se il primo operando non è compreso nell'intervallo.

74.4.3 Le relazioni dal punto di vista di SQL

SQL tratta le relazioni attraverso il modello tabellare; di conseguenza si adegua tutta la sua filosofia e il modo di esprimere i concetti nella sua documentazione. Pertanto, le relazioni per SQL sono delle «tabelle», che vengono definite nel modo seguente dalla documentazione standard.

- La tabella è un insieme di più righe. Una riga è una sequenza non vuota di valori. Ogni riga della stessa tabella ha la stessa cardinalità e contiene un valore per ogni colonna di quella tabella. L'i-esimo valore di ogni riga di una tabella è un valore dell'i-esima colonna di quella tabella. La riga è l'elemento che costituisce la più piccola unità di dati che può essere inserita in una tabella e cancellata da una tabella.
- Il grado di una tabella è il numero di colonne della stessa. In ogni momento, il grado della tabella è lo stesso della cardinalità di ognuna delle sue righe; la cardinalità della tabella (cioè il numero delle righe contenute) è la stessa della cardinalità di ognuna delle sue colonne. Una tabella la cui cardinalità sia zero viene definita come vuota.

In pratica, la tabella è un contenitore di informazioni organizzato in righe e colonne. La tabella viene identificata per nome, così anche le colonne, mentre le righe vengono identificate attraverso il loro contenuto.

Nel modello di SQL, le colonne sono ordinate, anche se ciò non è sempre un elemento indispensabile, dal momento che si possono identificare per nome. Inoltre sono ammissibili tabelle contenenti righe duplicate.

Si osservi, comunque, che nel resto di questo documento si preferisce la terminologia generale delle basi di dati, dove, al posto di tabelle, righe e colonne, si parla piuttosto di relazioni, tuple e attributi.

74.4.3.1 Creazione di una relazione

La creazione di una relazione avviene attraverso un'istruzione che può assumere un'articolazione molto complessa, a seconda delle caratteristiche particolari che da questa relazione si vogliono ottenere. La sintassi più semplice è quella seguente:

```
CREATE TABLE nome_relazione ( specifiche )
```

Tuttavia, sono proprio le specifiche indicate tra le parentesi tonde che possono tradursi in un sistema molto confuso. La creazione di una relazione elementare può essere espressa con la sintassi seguente:

```
CREATE TABLE nome_relazione (nome_attributo tipo [, ...])
```

In questo modo, all'interno delle parentesi vengono semplicemente elencati i nomi degli attributi seguiti dal tipo di dati che in essi possono essere contenuti. L'esempio seguente rappresenta l'istruzione necessaria a creare una relazione composta da cinque attributi, contenenti rispettivamente informazioni su: codice, cognome, nome, indirizzo e numero di telefono.

```
CREATE TABLE Indirizzi (
        Codice          integer,
        Cognome         char(40),
        Nome            char(40),
        Indirizzo       varchar(60),
        Telefono        varchar(40)
    )
```

74.4.3.2 Valori predefiniti

Quando si inseriscono delle tuple all'interno della relazione, in linea di principio è possibile che i valori corrispondenti ad attributi particolari non siano inseriti esplicitamente. Se si verifica questa situazione (purché ciò sia consentito dai vincoli), viene assegnato a questi elementi mancanti un valore predefinito. Questo può essere stabilito all'interno delle specifiche di creazione della relazione; in mancanza di tale definizione, viene assegnato '**NULL**', corrispondente al valore indefinito.

La sintassi necessaria a creare una relazione contenente le indicazioni sui valori predefiniti da utilizzare è la seguente:

```
CREATE TABLE nome_relazione (
                    nome_attributo tipo
                        [DEFAULT espressione]
                    [,…]
        )
```

L'esempio seguente crea la stessa relazione già vista nell'esempio precedente, specificando come valore predefinito per l'indirizzo, la stringa di caratteri: «sconosciuto».

```
CREATE TABLE Indirizzi (
        Codice          integer,
        Cognome         char(40),
        Nome            char(40),
        Indirizzo       varchar(60)   DEFAULT 'sconosciuto',
        Telefono        varchar(40)
    )
```

74.4.3.3 Vincoli interni alla relazione

«

Può darsi che in certe situazioni, determinati valori all'interno di una tupla non siano ammissibili, a seconda del contesto a cui si riferisce la relazione. I vincoli interni alla relazione sono quelli che possono essere risolti senza conoscere informazioni esterne alla relazione stessa.

Il vincolo più semplice da esprimere è quello di non ammissibilità dei valori indefiniti. La sintassi seguente ne mostra il modo.

```
CREATE TABLE nome_relazione (
                    nome_attributo  tipo
                        [NOT NULL]
                    [,…]
        )
```

L'esempio seguente crea la stessa relazione già vista negli esempi precedenti, specificando che il codice, il cognome, il nome e il telefono non possono essere indeterminati.

```
CREATE TABLE Indirizzi (
        Codice          integer         NOT NULL,
        Cognome         char(40)        NOT NULL,
        Nome            char(40)        NOT NULL,
        Indirizzo       varchar(60)     DEFAULT 'sconosciuto',
        Telefono        varchar(40)     NOT NULL
    )
```

Un altro vincolo importante è quello che permette di definire che un gruppo di attributi deve rappresentare dati unici in ogni tupla, cioè che non siano ammissibili tuple che per quel gruppo di attributi abbiano dati uguali. Segue lo schema sintattico relativo:

```
CREATE TABLE nome_relazione (
                nome_attributo tipo
                [,…],
                UNIQUE ( nome_attributo [,…] )
                [,…]
        )
```

L'indicazione dell'unicità può riguardare più gruppi di attributi in modo indipendente. Per ottenere questo si possono indicare più opzioni '**UNIQUE**'.

È il caso di osservare che il vincolo '**UNIQUE**' non è sufficiente per impedire che i dati possano essere indeterminati. Infatti, il valore indeterminato, '**NULL**', è diverso da ogni altro '**NULL**'.

L'esempio seguente crea la stessa relazione già vista negli esempi precedenti, specificando che i dati dell'attributo del codice devono

essere unici per ogni tupla.

```
CREATE TABLE Indirizzi (
        Codice          integer         NOT NULL,
        Cognome         char(40)        NOT NULL,
        Nome            char(40)        NOT NULL,
        Indirizzo       varchar(60)     DEFAULT 'sconosciuto',
        Telefono        varchar(40)     NOT NULL,
        UNIQUE (Codice)
    )
```

Quando un attributo, o un gruppo di attributi, costituisce un riferimento importante per identificare le varie tuple che compongono la relazione, si può utilizzare il vincolo '**`PRIMARY KEY`**', che può essere utilizzato una sola volta. Questo vincolo stabilisce anche che i dati contenuti, oltre a non poter essere doppi, non possono essere indefiniti.

```
CREATE TABLE nome_relazione (
                    nome_attributo tipo
                    [,...],
                    PRIMARY KEY ( nome_attributo [,...] )
        )
```

L'esempio seguente crea la stessa relazione già vista negli esempi precedenti specificando che l'attributo del codice deve essere considerato come chiave primaria.

```
CREATE TABLE Indirizzi (
        Codice          integer,
        Cognome         char(40)        NOT NULL,
        Nome            char(40)        NOT NULL,
        Indirizzo       varchar(60)     DEFAULT 'sconosciuto',
        Telefono        varchar(40)     NOT NULL,
        PRIMARY KEY (Codice)
    )
```

74.4.3.4 Vincoli esterni alla relazione

I vincoli esterni alla relazione riguardano principalmente la connessione con altre relazioni e la necessità che i riferimenti a queste siano validi. La definizione formale di questa connessione è molto complessa e qui non viene descritta. Si tratta, in ogni caso, dell'opzione '**`FOREIGN KEY`**' seguita da '**`REFERENCES`**'.

Vale la pena però di considerare i meccanismi che sono coinvolti. Infatti, nel momento in cui si inserisce un valore, il sistema può impedire l'operazione perché non valida in base all'assenza di quel valore in un'altra relazione esterna specificata. Il problema nasce però nel momento in cui nella relazione esterna viene eliminata o modificata una tupla che è oggetto di un riferimento da parte della prima. Si pongono le alternative seguenti.

Vincolo	Descrizione
`CASCADE`	Se nella relazione esterna il dato a cui si fa riferimento è stato cambiato, viene cambiato anche il riferimento nella relazione di partenza; se nella relazione esterna la tupla corrispondente viene rimossa, viene rimossa anche la tupla della relazione di partenza.
`SET NULL`	Se viene a mancare l'oggetto a cui si fa riferimento, viene modificato il dato attribuendo il valore indefinito.

Vincolo	Descrizione
`SET DEFAULT`	Se viene a mancare l'oggetto a cui si fa riferimento, viene modificato il dato attribuendo il valore predefinito.
`NO ACTION`	Se viene a mancare l'oggetto a cui si fa riferimento, non viene modificato il dato contenuto nella relazione di partenza.

Le azioni da compiere si possono distinguere in base all'evento che ha causato la rottura del riferimento: cancellazione della tupla della relazione esterna o modifica del suo contenuto.

74.4.3.5 Modifica della struttura della relazione

La modifica della struttura di una relazione riguarda principalmente la sua organizzazione in attributi. Le cose più semplici che si possono desiderare di fare sono l'aggiunta di nuovi attributi e l'eliminazione di attributi già esistenti. Vedendo il problema in questa ottica, la sintassi si riduce ai due casi seguenti:

```
ALTER TABLE nome_relazione ADD [COLUMN]
            nome_attributo tipo [altre_caratteristiche]
```

```
ALTER TABLE nome_relazione DROP [COLUMN] nome_attributo
```

Nel primo caso si aggiunge un attributo, del quale si deve specificare il nome, il tipo ed eventualmente i vincoli; nel secondo si tratta solo di indicare l'attributo da eliminare. A livello di singolo attributo può essere eliminato o assegnato un valore predefinito.

```
ALTER TABLE nome_relazione ALTER [COLUMN]
            nome_attributo DROP DEFAULT
```

```
ALTER TABLE nome_relazione ALTER [COLUMN]
            nome_attributo SET DEFAULT valore_predefinito
```

74.4.3.6 Eliminazione di una relazione

L'eliminazione di una relazione, con tutto il suo contenuto, è un'operazione semplice che dovrebbe essere autorizzata solo all'utente che l'ha creata.

```
DROP TABLE nome_relazione
```

74.5 DML

DML, ovvero *Data manipulation language*, è il linguaggio usato per inserire, modificare e accedere ai dati. In questa sezione viene trattato il linguaggio SQL per ciò che riguarda specificatamente l'inserimento, la lettura e la modifica del contenuto delle relazioni.

74.5.1 Inserimento, eliminazione e modifica dei dati

L'inserimento, l'eliminazione e la modifica dei dati di una relazione è un'operazione che interviene sempre a livello delle tuple. Infatti, come già definito, la tupla è l'elemento che costituisce l'unità di dati più piccola che può essere inserita o cancellata da una relazione.

74.5.1.1 Inserimento di tuple

«

L'inserimento di una nuova tupla all'interno di una relazione viene eseguito attraverso l'istruzione '**INSERT**'. Dal momento che nel modello di SQL gli attributi sono ordinati, è sufficiente indicare ordinatamente l'elenco dei valori della tupla da inserire, come mostra la sintassi seguente:

```
INSERT INTO nome_relazione VALUES (espressione_1 [,...espressione_n])
```

Per esempio, l'inserimento di una tupla nella relazione '**Indirizzi**' già mostrata in precedenza, potrebbe avvenire nel modo seguente:

```
INSERT INTO Indirizzi
    VALUES (
        01,
        'Pallino',
        'Pinco',
        'Via Biglie 1',
        '0222,222222'
    )
```

Se i valori inseriti sono meno del numero degli attributi della relazione, i valori mancanti, in coda, ottengono quanto stabilito come valore predefinito, o '**NULL**' in sua mancanza (sempre che ciò sia concesso dai vincoli della relazione).

L'inserimento dei dati può avvenire in modo più chiaro e sicuro elencando prima i nomi degli attributi, in modo da evitare di dipendere dalla sequenza degli attributi memorizzata nella relazione. La sintassi seguente mostra il modo di ottenere questo.

```
INSERT INTO nome_relazione (attributo_1 [,...attributo_n]) ]
    VALUES (espressione_1 [,...espressione_n] )
```

L'esempio già visto potrebbe essere tradotto nel modo seguente, più prolisso, ma anche più chiaro:

```
INSERT INTO Indirizzi (
        Codice,
        Cognome,
        Nome,
        Indirizzo,
        Telefono
    )
    VALUES (
        01,
        'Pallino',
        'Pinco',
        'Via Biglie 1',
        '0222,222222'
    )
```

Questo modo esplicito di fare riferimento agli attributi garantisce anche che eventuali modifiche di lieve entità nella struttura della relazione non debbano necessariamente riflettersi nei programmi. L'esempio seguente mostra l'inserimento di alcuni degli attributi della tupla, lasciando che gli altri ottengano l'assegnamento di un valore predefinito.

```
INSERT INTO Indirizzi (
        Codice,
        Cognome,
        Nome,
        Telefono
```

```
    )
    VALUES (
        01,
        'Pinco',
        'Pallino',
        '0222,222222'
    )
```

74.5.1.2 Aggiornamento delle tuple

La modifica delle tuple può avvenire attraverso una scansione della relazione, dalla prima all'ultima tupla, eventualmente controllando la modifica in base all'avverarsi di determinate condizioni. La sintassi per ottenere questo risultato, leggermente semplificata, è la seguente:

```
UPDATE relazione
    SET attributo_1=espressione_1 [ ,...attributo_n=espressione_n ]
    [WHERE condizione ]
```

L'istruzione '**UPDATE**' esegue tutte le sostituzioni indicate dalle coppie ***attributo=espressione***, per tutte le tuple in cui la condizione posta dopo la parola chiave '**WHERE**' si avvera. Se tale condizione manca, l'effetto delle modifiche si riflette su tutte le tuple della relazione.

L'esempio seguente aggiunge un attributo alla relazione degli indirizzi, per contenere il nome del comune di residenza; successivamente viene inserito il nome del comune «Sferopoli» in base al prefisso telefonico.

```
ALTER TABLE Indirizzi ADD COLUMN Comune char(30)
```

```
UPDATE Indirizzi
    SET Comune='Sferopoli'
    WHERE Telefono >= '022' AND Telefono < '023'
```

Eventualmente, al posto dell'espressione si può indicare la parola chiave '**DEFAULT**' che fa in modo di assegnare il valore predefinito per quel attributo.

74.5.1.3 Eliminazione di tuple

La cancellazione di tuple da una relazione è un'operazione molto semplice. Richiede solo l'indicazione del nome della relazione e la condizione in base alla quale le tuple devono essere cancellate.

```
DELETE FROM relazione [WHERE condizione]
```

Se la condizione non viene indicata, si cancellano tutte le tuple!

74.5.2 Interrogazioni di relazioni

L'interrogazione di una relazione è l'operazione con cui si ottengono i dati contenuti al suo interno, in base a dei criteri di filtro determinati. L'interrogazione consente anche di combinare assieme dati provenienti da relazioni differenti, in base a dei «collegamenti» che possono intercorrere tra queste.

74.5.2.1 Interrogazioni elementari

La forma più semplice di esprimere la sintassi necessaria a interrogare **una** sola relazione è quella espressa dallo schema seguente:

```
SELECT espress_col_1 [ , ...espress_col_n ]
    FROM relazione
    [WHERE condizione ]
```

In questo modo è possibile definire gli attributi che si intendono utilizzare per il risultato, mentre le tuple si specificano, eventualmente, con la condizione posta dopo la parola chiave '**WHERE**'. L'esempio seguente mostra la proiezione degli attributi del cognome e nome della relazione di indirizzi già vista negli esempi delle altre sezioni, senza porre limiti alle tuple.

```
SELECT Cognome, Nome FROM Indirizzi
```

Quando si vuole ottenere una selezione composta dagli stessi attributi della relazione originale, nel suo stesso ordine, si può utilizzare un carattere jolly particolare, l'asterisco ('*'). Questo rappresenta l'elenco di tutti gli attributi della relazione indicata.

```
SELECT * FROM Indirizzi
```

È bene osservare che gli attributi si esprimono attraverso un'espressione, questo significa che gli attributi a cui si fa riferimento sono quelle del risultato finale, cioè della relazione che viene restituita come selezione o proiezione della relazione originale. L'esempio seguente emette un solo attributo contenente un ipotetico prezzo scontato del 10 %, in pratica viene moltiplicato il valore di un attributo contenente il prezzo per 0,90, in modo da ottenerne il 90 % (100 % meno lo sconto).

```
SELECT Prezzo * 0.90 FROM Listino
```

In questo senso si può comprendere l'utilità di assegnare esplicita-

mente un nome agli attributi del risultato finale, come indicato dalla sintassi seguente:

```
SELECT espress_col_1 AS nome_col_1 [,...espress_col_n AS nome_col_n]
    FROM relazione
    [WHERE condizione]
```

In questo modo, l'esempio precedente può essere trasformato come segue, dando un nome all'attributo generato e chiarendone così il contenuto.

```
SELECT Prezzo * 0.90 AS Prezzo_Scontato FROM Listino
```

Finora è stata volutamente ignorata la condizione che controlla le tuple da selezionare. Anche se potrebbe essere evidente, è bene chiarire che la condizione posta dopo la parola chiave '**WHERE**' può fare riferimento solo ai dati originali della relazione da cui si attingono. Quindi, non è valida una condizione che utilizza un riferimento a un nome che appare dopo la parola chiave '**AS**' abbinata alle espressioni degli attributi.

Per qualche motivo che viene chiarito in seguito, può essere conveniente associare un alias alla relazione da cui estrarre i dati. Anche in questo caso si utilizza la parola chiave '**AS**', come indicato dalla sintassi seguente:

```
SELECT specificazione_dell'attributo_1 [,...specificazione_dell'attributo_n]
    FROM relazione AS alias
    [WHERE condizione]
```

Quando si vuole fare riferimento al nome di un attributo, se per qual-

che motivo questo nome dovesse risultare ambiguo, si può aggiungere anteriormente il nome della relazione a cui appartiene, separandolo attraverso l'operatore punto ('.'). L'esempio seguente è la proiezione dei cognomi e dei nomi della solita relazione degli indirizzi. In questo caso, le espressioni degli attributi rappresentano solo gli attributi corrispondenti della relazione originaria, con l'aggiunta dell'indicazione esplicita del nome della relazione stessa.

```
SELECT Indirizzi.Cognome, Indirizzi.Nome FROM Indirizzi
```

A questo punto, se al nome della relazione viene abbinato un alias, si può esprimere la stessa cosa indicando il nome dell'alias al posto di quello della relazione, come nell'esempio seguente:

```
SELECT Ind.Cognome, Ind.Nome FROM Indirizzi AS Ind
```

74.5.2.2 Interrogazioni ordinate

«

Per ottenere un elenco ordinato in base a qualche criterio, si utilizza l'istruzione '**SELECT**' con l'indicazione di un'espressione in base alla quale effettuare l'ordinamento. Questa espressione è preceduta dalle parole chiave '**ORDER BY**':

```
SELECT espress_col_1 [,...espress_col_n]
    FROM relazione
    [WHERE condizione]
    ORDER BY espressione [ASC|DESC] [,...]
```

L'espressione può essere il nome di un attributo, oppure un'espressione che genera un risultato da uno o più attributi; l'aggiunta eventuale della parola chiave '**ASC**', o '**DESC**', permette di specificare un

ordinamento crescente, o discendente. Come si vede, le espressioni di ordinamento possono essere più di una, separate con una virgola.

```
SELECT Cognome, Nome FROM Indirizzi ORDER BY Cognome
```

L'esempio mostra un'applicazione molto semplice del problema, in cui si ottiene un elenco dei soli attributi '**Cognome**' e '**Nome**', della relazione '**Indirizzi**', ordinato per '**Cognome**'.

```
SELECT Cognome, Nome FROM Indirizzi ORDER BY Cognome, Nome
```

Questo esempio, aggiunge l'indicazione del nome nella chiave di ordinamento, in modo che in presenza di cognomi uguali, la scelta venga fatta in base al nome.

```
SELECT Cognome, Nome
    FROM Indirizzi
    ORDER BY TRIM( Cognome ), TRIM( Nome )
```

Questo ultimo esempio mostra l'utilizzo di due espressioni come chiave di ordinamento. Per la precisione, la funzione ***TRIM()***, usata in questo modo, serve a eliminare gli spazi iniziali e finali superflui. In questo modo, se i nomi e i cognomi sono stati inseriti con degli spazi iniziali, questi non vanno a influire sull'ordinamento.

74.5.2.3 Interrogazioni simultanee di più relazioni

Se dopo la parola chiave '**FROM**' si indicano più relazioni (ciò vale anche se si indica più volte la stessa relazione), si intende fare riferimento a una relazione generata dal prodotto di queste. Se per esempio si vogliono abbinare due relazioni, una di tre tuple con due attributi e un'altra di due tuple con due attributi, quello che si ottiene è una relazione con quattro attributi composta da sei tuple. Infatti,

ogni tupla della prima relazione risulta abbinata con ogni tupla della seconda.

```
SELECT specificazione_dell'attributo_1 [ , ...specificazione_dell'attributo_n ]
    FROM specificazione_della_relazione_1 [ , ...specificazione_della_relazione_n ]
    [WHERE condizione ]
```

Viene proposto un esempio banalizzato, con il quale poi si vuole eseguire un'elaborazione (figura 74.54).

Figura 74.54. Relazioni '**Articoli**' e '**Movimenti**' di una gestione del magazzino ipotetica.

Articoli	
Codice	**Descrizione**
vite30	Vite 3 mm
dado30	Dado 3 mm
rond50	Rondella 5 mm

Movimenti			
Codice	**Data**	**Carico**	**Scarico**
dado30	01/01/2012	1200	
vite30	01/01/2012		800
vite30	03/01/2012	2000	
rond50	03/01/2012		500

Da questa situazione si vuole ottenere la congiunzione della relazione '**Movimenti**' con tutte le informazioni corrispondenti della relazione '**Articoli**', basando il riferimento sull'attributo '**Codice**'. In pratica si vuole ottenere la relazione della figura 74.55.

Tabella 74.55. Risultato del join che si intende ottenere tra la relazione '**Movimenti**' e la relazione '**Articoli**'.

Codice	Data	Carico	Scarico	Descrizione
dado30	01/01/2012	1200		Dado 3 mm
vite30	01/01/2012	2000		Vite 3 mm
vite30	03/01/2012		800	Vite 3 mm
rond50	03/01/2012		500	Rondella 5 mm

Considerato che da un'istruzione '**SELECT**' contenente il riferimento a più relazioni si genera il prodotto tra queste, si pone poi il problema di eseguire una proiezione degli attributi desiderati e, soprattutto, di selezionare le tuple. In questo caso, la selezione deve essere basata sulla corrispondenza tra l'attributo '**Codice**' della prima relazione, con lo stesso attributo della seconda. Dovendo fare riferimento a due attributi di relazioni differenti, aventi però lo stesso nome, diviene indispensabile indicare i nomi degli attributi prefissandoli con i nomi delle relazioni rispettive.

```
SELECT
    Movimenti.Codice,
    Movimenti.Data,
    Movimenti.Carico,
    Movimenti.Scarico,
    Articoli.Descrizione
    FROM Movimenti, Articoli
    WHERE Movimenti.Codice = Articoli.Codice;
```

L'interrogazione simultanea di più relazioni si presta anche per elaborazioni della stessa relazione più volte. In tal caso, diventa obbligatorio l'uso degli alias. Si osservi il caso seguente:

```
SELECT Ind1.Cognome, Ind1.Nome
    FROM Indirizzi AS Ind1, Indirizzi AS Ind2
    WHERE
        Ind1.Cognome = Ind2.Cognome
    AND Ind1.Nome <> Ind2.Nome
```

Il senso di questa interrogazione, che utilizza la stessa relazione degli indirizzi per due volte con due alias differenti, è quello di ottenere l'elenco delle persone che hanno lo stesso cognome, avendo però un nome differente.

Esiste anche un'altra situazione in cui si ottiene l'interrogazione simultanea di più relazioni: l'***unione***. Si tratta semplicemente di attaccare il risultato di un'interrogazione su una relazione con quello di un'altra relazione, quando gli attributi finali appartengono allo stesso tipo di dati.

```
SELECT specificazione_attributo_1 [ , ...specificazione_attributo_n ]
        FROM specificazione_relazione_1 [ , ...specificazione_relazione_n ]
        [WHERE condizione ]
    UNION
        SELECT specificatore_attributo_1 [ , ...specificazione_attributo_n ]
            FROM specificazione_relazione_1 [ , ...specificazione_relazione_n ]
            [WHERE condizione ]
```

Lo schema sintattico dovrebbe essere abbastanza esplicito: si uniscono due istruzioni '**SELECT**' in un risultato unico, attraverso la parola chiave '**UNION**'.

74.5.2.4 Condizioni

La condizione che esprime la selezione delle tuple può essere composta come si vuole, purché il risultato sia di tipo logico e i dati a cui si fa riferimento provengano dalle relazioni di partenza. Quindi si possono usare anche altri operatori di confronto, funzioni e operatori booleani.

È bene ricordare che il valore indefinito, rappresentato da '**NULL**', è diverso da qualunque altro valore, compreso un altro valore indefinito. Per verificare che un valore sia o non sia indefinito, si deve usare l'operatore '**IS NULL**' oppure '**IS NOT NULL**'.

74.5.2.5 Aggregazioni

L'aggregazione è una forma di interrogazione attraverso cui si ottengono risultati riepilogativi del contenuto di una relazione, in forma di relazione contenente una sola tupla. Per questo si utilizzano delle funzioni speciali al posto dell'espressione che esprime gli attributi del risultato. Queste funzioni restituiscono un solo valore e come tali concorrono a creare un'unica tupla. Le funzioni di aggregazione sono: ***COUNT()***, ***SUM()***, ***MAX()***, ***MIN()***, ***AVG()***. Per intendere il problema, si osservi l'esempio seguente:

```
SELECT COUNT(*) FROM Movimenti WHERE …
```

In questo caso, quello che si ottiene è solo il numero di tuple della relazione '**Movimenti**' che soddisfano la condizione posta dopo la parola chiave '**WHERE**' (qui non è stata indicata). L'asterisco posto co-

me parametro della funzione ***COUNT()*** rappresenta effettivamente l'elenco di tutti i nomi degli attributi della relazione '**Movimenti**'.

Quando si utilizzano funzioni di questo tipo, occorre considerare che l'elaborazione si riferisce alla relazione virtuale generata dopo la selezione posta da '**WHERE**'.

La funzione ***COUNT()*** può essere descritta attraverso la sintassi seguente:

```
COUNT( * )
```

```
COUNT( [DISTINCT|ALL] lista_attributi)
```

Utilizzando la forma già vista, quella dell'asterisco, si ottiene solo il numero delle tuple della relazione. L'opzione '**DISTINCT**', seguita da una lista di nomi di attributi, fa in modo che vengano contate le tuple contenenti valori differenti per quel gruppo di attributi. L'opzione '**ALL**' è implicita quando non si usa '**DISTINCT**' e indica semplicemente di contare tutte le tuple.

Il conteggio delle tuple esclude in ogni caso quelle in cui il contenuto di tutti gli attributi selezionati è indefinito ('**NULL**').

Le altre funzioni aggreganti non prevedono l'asterisco, perché fanno riferimento a un'espressione che genera un risultato per ogni tupla ottenuta dalla selezione.

```
SUM( [DISTINCT|ALL] espressione)
```

```
MAX( [DISTINCT|ALL] espressione)
```

```
MIN( [DISTINCT|ALL] espressione)
```

```
AVG( [DISTINCT|ALL] espressione)
```

In linea di massima, per tutti questi tipi di funzioni aggreganti, l'espressione deve generare un risultato numerico, sul quale calcolare la sommatoria, ***SUM()***, il valore massimo, ***MAX()***, il valore minimo, ***MIN()***, la media, ***AVG()***.

L'esempio seguente calcola lo stipendio medio degli impiegati, ottenendo i dati da un'ipotetica relazione '**`Emolumenti`**', limitandosi ad analizzare le tuple riferite a un certo settore.

```
SELECT AVG( Stipendio ) FROM Emolumenti
    WHERE Settore = 'Amministrazione'
```

L'esempio seguente è una variante in cui si estraggono rispettivamente lo stipendio massimo, medio e minimo.

```
SELECT MAX( Stipendio ), AVG( Stipendio ), MIN( Stipendio )
    FROM Emolumenti WHERE Settore = 'Amministrazione'
```

L'esempio seguente è invece volutamente **errato**, perché si mescolano funzioni aggreganti assieme a espressioni di attributi normali.

```
-- Esempio errato
SELECT MAX( Stipendio ), Settore FROM Emolumenti
    WHERE Settore = 'Amministrazione'
```

74.5.2.6 Raggruppamenti

«

Le aggregazioni possono essere effettuate in riferimento a gruppi di tuple, distinguibili in base al contenuto di uno o più attributi. In questo tipo di interrogazione si può generare solo una relazione composta da tanti attributi quanti sono quelli presi in considerazione dall'opzione di raggruppamento, assieme ad altre contenenti solo espressioni di aggregazione.

Alla sintassi normale già vista nelle sezioni precedenti, si aggiunge la clausola '**GROUP BY**'.

```
SELECT specificazione_dell'attributo_1 [ , ...specificazione_dell'attributo_n ]
    FROM specificazione_della_relazione_1 [ , ...specificazione_della_relazione_n ]
    [WHERE condizione ]
    GROUP BY attributo_1 [ , ... ]
```

Per comprendere l'effetto di questa sintassi, si deve scomporre idealmente l'operazione di selezione da quella di raggruppamento:

1. la relazione ottenuta dall'istruzione '**SELECT**...**FROM**' viene filtrata dalla condizione '**WHERE**';
2. la relazione risultante viene riordinata in modo da raggruppare le tuple in cui i contenuti degli attributi elencati dopo l'opzione '**GROUP BY**' sono uguali;
3. su questi gruppi di tuple vengono valutate le funzioni di aggregazione.

Figura 74.62. Carichi e scarichi in magazzino.

Movimenti				
Codice	**Data**	**Carico**	**Scarico**	**...**
vite40	01/01/2012	1200		...
vite30	01/01/2012		800	...
vite40	01/01/2012	1500		...
vite30	02/01/2012		1000	...
vite30	03/01/2012	2000		...
rond50	03/01/2012		500	...
vite40	04/01/2012	2200		...

Si osservi la relazione riportata in figura 74.62, mostra la solita sequenza di carichi e scarichi di magazzino. Si potrebbe porre il problema di conoscere il totale dei carichi e degli scarichi per ogni articolo di magazzino. La richiesta può essere espressa con l'istruzione seguente:

```
SELECT Codice, SUM( Carico ), SUM( Scarico ) FROM Movimenti
    GROUP BY Codice
```

Quello che si ottiene appare nella figura 74.64.

Figura 74.64. Carichi e scarichi totali.

Codice	**SUM(Carico)**	**SUM(Scarico)**
vite40	4900	
vite30	2000	1800
rond50		500

Volendo si possono fare i raggruppamenti in modo da avere i totali distinti anche in base al giorno, come nell'istruzione seguente:

```
SELECT Codice, Data, SUM( Carico ), SUM( Scarico )
    FROM Movimenti GROUP BY Codice, Data
```

Come già affermato, la condizione posta dopo la parola chiave '**WHERE**' serve a filtrare inizialmente le tuple da considerare nel raggruppamento. Se quello che si vuole è filtrare ulteriormente il risultato di un raggruppamento, occorre usare la clausola '**HAVING**'.

```
SELECT specificazione_dell'attributo_1 [, ...specificazione_dell'attributo_n]
    FROM specificazione_della_relazione_1 [, ...specificazione_della_relazione_n]
    [WHERE condizione]
    GROUP BY attributo_1 [, ...]
    HAVING condizione
```

L'esempio seguente serve a ottenere il raggruppamento dei carichi e scarichi degli articoli, limitando però il risultato a quelli per i quali sia stata fatta una quantità di scarichi consistente (superiore a 1000 unità).

```
SELECT Codice, SUM( Carico ), SUM( Scarico ) FROM Movimenti
    GROUP BY Codice
    HAVING SUM( Scarico ) > 1000
```

Dall'esempio già visto in figura 74.64 risulterebbe escluso l'articolo '**rond50**'.

74.5.3 Trasferimento di dati in un'altra relazione

«

Alcune forme particolari di interrogazioni SQL possono essere utilizzate per inserire dati in relazioni esistenti o per crearne di nuove.

74.5.3.1 Creazione di una nuova relazione a partire da altre

L'istruzione '**SELECT**' può servire per creare una nuova relazione a partire dai dati ottenuti dalla sua interrogazione.

```
SELECT specificazione_dell'attributo_1 [ , ...specificazione_dell'attributo_n ]
    INTO TABLE relazione_da_generare
    FROM specificazione_della_relazione_1 [ , ...specificazione_della_relazione_n ]
    [WHERE condizione ]
```

L'esempio seguente crea la relazione '**Mia_prova**' come risultato della fusione delle relazioni '**Indirizzi**' e '**Presenze**'.

```
SELECT
    Presenze.Giorno,
    Presenze.Ingresso,
    Presenze.Uscita,
    Indirizzi.Cognome,
    Indirizzi.Nome
    INTO TABLE Mia_prova
    FROM Presenze, Indirizzi
    WHERE Presenze.Codice = Indirizzi.Codice;
```

74.5.3.2 Inserimento in una relazione esistente

L'inserimento di dati in una relazione esistente prelevando da dati contenuti in altre, può essere fatta attraverso l'istruzione '**INSERT**' sostituendo la clausola '**VALUES**' con un'interrogazione ('**SELECT**').

```
INSERT INTO nome_relazione [(attributo_1...attributo_n)]
    SELECT espressione_1, ... espressione_n
        FROM relazioni_di_origine
        [WHERE condizione]
```

L'esempio seguente aggiunge alla relazione dello storico delle presenze le registrazioni vecchie che poi vengono cancellate.

```
INSERT INTO PresenzeStorico (
        PresenzeStorico.Codice,
        PresenzeStorico.Giorno,
        PresenzeStorico.Ingresso,
        PresenzeStorico.Uscita
    )
    SELECT
        Presenze.Codice,
        Presenze.Giorno,
        Presenze.Ingresso,
        Presenze.Uscita
        FROM Presenze
        WHERE Presenze.Giorno <= '01/01/1999';

DELETE FROM Presenze WHERE Giorno <= '01/01/1999';
```

74.5.4 Viste

«

Le viste sono delle relazioni virtuali ottenute a partire da relazioni vere e proprie o da altre viste, purché non si formino ricorsioni. Il concetto non dovrebbe risultare strano. In effetti, il risultato delle interrogazioni è sempre in forma di relazione. La vista crea una sorta di interrogazione permanente che acquista la personalità di una relazione normale.

```
CREATE VIEW nome_vista [(attributo_1 [,...attributo_n)]]
    AS richiesta
```

Dopo la parola chiave '**AS**' deve essere indicato ciò che compone un'istruzione '**SELECT**'. L'esempio seguente, genera la vista dei movimenti di magazzino del solo articolo '**vite30**'.

```
CREATE VIEW Movimenti_Vite30
    AS SELECT Codice, Data, Carico, Scarico
        FROM Movimenti
        WHERE Codice = 'vite30'
```

L'eliminazione di una vista si ottiene con l'istruzione '**DROP VIEW**', come illustrato dallo schema sintattico seguente:

```
DROP VIEW nome_vista
```

Volendo eliminare la vista '**Movimenti_Vite30**', si può intervenire semplicemente come nell'esempio seguente:

```
DROP VIEW Movimenti_Vite30
```

74.5.5 Cursori

Quando il risultato di un'interrogazione SQL deve essere gestito all'interno di un programma, si pone un problema nel momento in cui ciò che si ottiene è composto da più di una sola tupla. Per poter scorrere un elenco ottenuto attraverso un'istruzione '**SELECT**', tupla per tupla, si deve usare un ***cursore***.

La dichiarazione e l'utilizzo di un cursore avviene all'interno di una transazione. Quando la transazione si chiude attraverso un'istruzione '**COMMIT**' o '**ROLLBACK**', si chiude anche il cursore.

74.5.5.1 Dichiarazione e apertura

«

L'SQL prevede due fasi prima dell'utilizzo di un cursore: la dichiarazione e la sua apertura:

```
DECLARE cursore [INSENSITIVE] [SCROLL] CURSOR FOR
    SELECT …
```

```
OPEN cursore
```

Nella dichiarazione, la parola chiave '**INSENSITIVE**' serve a stabilire che il risultato dell'interrogazione che si scandisce attraverso il cursore, non deve essere sensibile alle variazioni dei dati originali; la parola chiave '**SCROLL**' indica che è possibile estrarre più tuple simultaneamente attraverso il cursore.

```
DECLARE Mio_cursore CURSOR FOR
    SELECT
        Presenze.Giorno,
        Presenze.Ingresso,
        Presenze.Uscita,
        Indirizzi.Cognome,
        Indirizzi.Nome
        FROM Presenze, Indirizzi
        WHERE Presenze.Codice = Indirizzi.Codice;
```

L'esempio mostra la dichiarazione del cursore '**Mio_cursore**', abbinato alla selezione degli attributi composti dal collegamento di due relazioni, '**Presenze**' e '**Indirizzi**', dove le tuple devono avere lo stesso numero di codice. Per attivare questo cursore, lo si deve aprire come nell'esempio seguente:

```
OPEN Mio_cursore
```

74.5.5.2 Scansione

La scansione di un'interrogazione inserita in un cursore, avviene attraverso l'istruzione '**FETCH**'. Il suo scopo è quello di estrarre una tupla alla volta, in base a una posizione, relativa o assoluta.

```
FETCH [ [ NEXT | PRIOR | FIRST | LAST
              | { ABSOLUTE | RELATIVE } n ]
    FROM cursore ]
    INTO :variabile [,…]
```

Le parole chiave '**NEXT**', '**PRIOR**', '**FIRST**', '**LAST**', permettono rispettivamente di ottenere la tupla successiva, quella precedente, la prima e l'ultima. Le parole chiave '**ABSOLUTE**' e '**RELATIVE**' sono seguite da un numero, corrispondente alla scelta della tupla ***n***-esima, rispetto all'inizio del gruppo per il quale è stato definito il cursore ('**ABSOLUTE**'), oppure della tupla ***n***-esima rispetto all'ultima tupla estratta da un'istruzione '**FETCH**' precedente.

Le variabili indicate dopo la parola chiave '**INTO**', che in particolare sono precedute da due punti ('**:**'), ricevono ordinatamente il contenuto dei vari attributi della tupla estratta. Naturalmente, le variabili in questione devono appartenere a un linguaggio di programmazione che incorpora l'SQL, dal momento che l'SQL stesso non fornisce questa possibilità.

```
FETCH NEXT FROM Mio_cursore
```

L'esempio mostra l'uso tipico di questa istruzione, dove si legge la tupla successiva (se non ne sono state lette fino a questo punto, si

tratta della prima), dal cursore dichiarato e aperto precedentemente. L'esempio seguente è identico dal punto di vista funzionale.

```
FETCH RELATIVE 1 FROM Mio_cursore
```

I due esempi successivi sono equivalenti e servono a ottenere la tupla precedente.

```
FETCH PRIOR FROM Mio_cursore
```

```
FETCH RELATIVE -1 FROM Mio_cursore
```

74.5.5.3 Chiusura

«

Il cursore, al termine dell'utilizzo, deve essere chiuso:

```
CLOSE cursore
```

Seguendo gli esempi visti in precedenza, per chiudere il cursore '**`Mio_cursore`**' basta l'istruzione seguente:

```
CLOSE Mio_cursore
```

74.6 DCL

«

DCL, ovvero *Data control language*, è il linguaggio usato per il «controllo» delle basi di dati. In questa sezione viene trattato il linguaggio SQL per ciò che riguarda la gestione delle basi di dati, degli utenti, dei privilegi assegnati loro e il controllo delle transazioni.

74.6.1 Gestione delle utenze

«

La gestione degli accessi in una base di dati è molto importante e potenzialmente indipendente dall'eventuale gestione degli utenti del sistema operativo sottostante. Per quanto riguarda i sistemi Unix, il

DBMS può riutilizzare la definizione degli utenti del sistema operativo, farvi riferimento, oppure astrarsi completamente (spesso vale questa ultima ipotesi).

Un DBMS SQL richiede la presenza di almeno un amministratore complessivo, che come tale abbia sempre tutti i privilegi necessari a intervenire come vuole nel DBMS. Il nome simbolico predefinito per questo utente dal linguaggio SQL standard è '**_SYSTEM**'.

Il sistema di definizione e controllo delle utenze è esterno al linguaggio SQL standard; tuttavia, i DBMS principali utilizzano istruzioni abbastanza uniformi per questo scopo.

Per la creazione di un utente si dispone normalmente dell'istruzione '**CREATE USER**', con opzioni che dipendono dalle caratteristiche particolari del DBMS, nella gestione delle utenze. I due modelli sintattici successivi si riferiscono, rispettivamente, a Oracle e a PostgreSQL, ma omettono varie opzioni specifiche e presumono che l'utente debba essere identificato attraverso una parola d'ordine:

```
CREATE USER nome_utente IDENTIFIED BY 'parola_d'ordine'
```

```
CREATE USER nome_utente [WITH PASSWORD 'parola_d'ordine']
```

Nel caso di MySQL, invece di introdurre un'istruzione che non esiste nello standard, si estende quella con cui si concedono i privilegi (descritta in un'altra sezione):

```
GRANT privilegi
    ON risorsa [, ...]
    TO utente
    IDENTIFIED BY 'parola_d'ordine'
    [WITH GRANT OPTION]
```

In tal modo, attribuendo dei privilegi a un utente, se questo non esiste ancora, viene creato contestualmente. Si osservi comunque, che le versioni più recenti di MySQL dispongono di un'istruzione '**CREATE USER**' simile a quella di altri DBMS.

Per l'eliminazione di un utente si dispone normalmente dell'istruzione '**DROP USER**', con opzioni che di solito consentono l'eliminazione contestuale di tutto ciò che appartiene a tale utente:

```
DROP USER nome_utente
```

Per modificare la parola d'ordine di un utente, si dispone normalmente dell'istruzione '**ALTER USER**', con la quale si potrebbero cambiare anche altre opzioni legate ai privilegi generali di cui può disporre tale utente. I due modelli sintattici successivi si riferiscono, rispettivamente, a Oracle e a PostgreSQL, omettendo opzioni che non sono indispensabili:

```
ALTER USER nome_utente IDENTIFIED BY 'parola_d'ordine'
```

```
ALTER USER nome_utente [WITH PASSWORD 'parola_d'ordine']
```

Nel caso di MySQL si usa una forma differente:

```
SET PASSWORD FOR nome_utente = PASSWORD('parola_d'ordine')
```

74.6.2 Gestione delle basi di dati

La creazione e l'eliminazione delle basi di dati è una funzione non considerata dallo standard SQL, anche se è normale che un DBMS consenta la gestione di più basi di dati simultaneamente. Pertanto, i vari DBMS offrono delle istruzioni SQL abbastanza uniformi:

```
CREATE DATABASE nome_base_di_dati
```

Di solito, salvo indicazione diversa derivante da opzioni particolari aggiunte all'istruzione, l'utente che crea la base di dati ne diviene il proprietario, con ogni facoltà sulla stessa, anche quella di eliminarla.

È il caso di osservare che uno dei problemi tecnici da considerare nella creazione di una base di dati sta nel definire la codifica da usare per la memorizzazione delle informazioni testuali. Di solito, questo genere di cose viene definito tramite delle opzioni specifiche, che si aggiungono al modello sintetico e generalizzato mostrato qui.

L'eliminazione di una base di dati richiede generalmente un'istruzione altrettanto semplice:

```
DROP DATABASE nome_base_di_dati
```

La differenza più importante tra i vari DBMS consiste nel modo di comportarsi di fronte a questo comando, quando la base di dati non è vuota. Se esiste un contenuto, la cancellazione potrebbe essere ri-

fiutata, oppure potrebbe essere ammessa se si aggiungono opzioni specifiche che servono a confermarne l'eliminazione. Tuttavia, non si può contare su un controllo di questo genere e la cancellazione di una base di dati richiede sempre la dovuta prudenza.

74.6.3 Gestione dei privilegi standard

«

L'utente che crea una relazione, o un'altra risorsa, è il suo creatore. Su tale risorsa è l'unico utente che possa modificarne la struttura e che possa eliminarla. In pratica è l'unico che possa usare le istruzioni '**`DROP`**' e '**`ALTER`**'. Chi crea una relazione, o un'altra risorsa, può concedere o revocare i privilegi degli altri utenti su di essa.

I privilegi che si possono concedere o revocare su una risorsa sono di vario tipo, espressi attraverso una parola chiave particolare. È bene considerare i casi seguenti:

Privilegio	Descrizione
`SELECT`	rappresenta l'operazione di lettura del valore di un oggetto della risorsa, per esempio dei valori di una tupla da una relazione (in pratica si riferisce all'uso dell'istruzione '**`SELECT`**');
`INSERT`	rappresenta l'azione di inserire un nuovo oggetto nella risorsa, come l'inserimento di una tupla in una relazione;
`UPDATE`	rappresenta l'operazione di aggiornamento del valore di un oggetto della risorsa, per esempio la modifica del contenuto di una tupla di una relazione;
`DELETE`	rappresenta l'eliminazione di un oggetto dalla risorsa, come la cancellazione di una tupla da una relazione;

Privilegio	Descrizione
`ALL PRIVILEGES`	rappresenta simultaneamente tutti i privilegi possibili riferiti a un oggetto.

I privilegi su una relazione, o su un'altra risorsa, vengono concessi attraverso l'istruzione '**GRANT**':

```
GRANT privilegi
    ON risorsa [,…]
    TO utenti
    [WITH GRANT OPTION]
```

Nella maggior parte dei casi, le risorse da controllare coincidono con una relazione. L'esempio seguente permette all'utente '**Pippo**' di leggere il contenuto della relazione '**Movimenti**':

```
GRANT SELECT ON Movimenti TO Pippo
```

L'esempio seguente, concede tutti i privilegi sulla relazione '**Movimenti**' agli utenti '**Pippo**' e '**Arturo**':

```
GRANT ALL PRIVILEGES ON Movimenti TO Pippo, Arturo
```

L'opzione '**WITH GRANT OPTION**' permette agli utenti presi in considerazione di concedere a loro volta tali privilegi ad altri utenti. L'esempio seguente concede all'utente '**Pippo**' di accedere in lettura al contenuto della relazione '**Movimenti**' e gli permette di concedere lo stesso privilegio ad altri:

```
GRANT SELECT ON Movimenti TO Pippo WITH GRANT OPTION
```

I privilegi su una relazione, o un'altra risorsa, vengono revocati attraverso l'istruzione '**REVOKE**':

```
REVOKE privilegi
    ON risorsa [, ...]
    FROM utenti
```

L'esempio seguente toglie all'utente '**Pippo**' il permesso di accedere in lettura al contenuto della relazione '**Movimenti**':

```
REVOKE SELECT ON Movimenti FROM Pippo
```

L'esempio seguente toglie tutti i privilegi sulla relazione '**Movimenti**' agli utenti '**Pippo**' e '**Arturo**':

```
REVOKE ALL PRIVILEGES ON Movimenti FROM Pippo, Arturo
```

74.6.4 Controllo delle transazioni

«

Una transazione SQL, è una sequenza di istruzioni che rappresenta un corpo unico dal punto di vista della memorizzazione effettiva dei dati. In altre parole, secondo l'SQL, la registrazione delle modifiche apportate alla base di dati avviene in modo asincrono, raggruppando assieme l'effetto di gruppi di istruzioni determinati.

Una transazione inizia nel momento in cui l'interprete SQL incontra, generalmente, l'istruzione '**START TRANSACTION**', terminando con l'istruzione '**COMMIT**', oppure '**ROLLBACK**': nel primo caso si conferma la transazione che viene memorizzata regolarmente, mentre nel secondo si richiede di annullare le modifiche apportate dalla transazione.

```
START TRANSACTION
```

```
COMMIT [WORK]
```

```
ROLLBACK [WORK]
```

Stando così le cose, si intende la necessità di utilizzare regolarmente l'istruzione '**COMMIT**' per memorizzare i dati quando non esiste più la necessità di annullare le modifiche.

```
START TRANSACTION
…
COMMIT

INSERT INTO Indirizzi
    VALUES (
        01,
        'Pallino',
        'Pinco',
        'Via Biglie 1',
        '0222,222222'
    )

COMMIT
```

L'esempio mostra un uso intensivo dell'istruzione '**COMMIT**', dove dopo l'inserimento di una tupla nella relazione '**Indirizzi**', viene confermata immediatamente la transazione.

```
START TRANSACTION
…
COMMIT

INSERT INTO Indirizzi
    VALUES (
        01,
        'Pallino',
        'Pinco',
        'Via Biglie 1',
        '0222,222222'
    )

ROLLBACK
```

Questo esempio mostra un ripensamento (per qualche motivo). Dopo l'inserimento di una tupla nella relazione '**Indirizzi**', viene annullata la transazione, riportando la relazione allo stato precedente.

74.7 Riferimenti

«

- Paolo Atzeni, Stefano Ceri, Stefano Paraboschi, Riccardo Torlone, *Basi di dati, concetti, linguaggi e architetture*, McGraw-Hill
- James Hoffmann, *Introduction to Structured Query Language*, *http://www.highcroft.com/highcroft/sql_intro.pdf*
- *JCC's SQL Std. Page*, *http://www.jcc.com/sql.htm*
- *SQL Reference Page*, *http://www.contrib.andrew.cmu.edu/~shadow/sql.html*

- *SQL92 BNF*, http://www.contrib.andrew.cmu.edu/~shadow/sql/sql2bnf.aug92.txt
- *ISO/IEC 9075:1992, Database Language SQL*, http://www.contrib.andrew.cmu.edu/~shadow/sql/sql1992.txt
- *BNF Grammar for ISO/IEC 9075:1999 - Database Language SQL (SQL-99)*, http://savage.net.au/SQL/sql-99.bnf , http://savage.net.au/SQL/sql-99.bnf.html
- *PostgreSQL*, http://www.postgresql.org/
- *MySQL*, http://www.mysql.com/
- *MaxDB*, http://www.mysql.com/sap/
- *Firebird*, http://firebird.sourceforge.net/