

Informatica Generale

Andrea Corradini

19 - Sistemi di Gestione delle Basi di Dati

Sommario

- Concetti base di Basi di Dati
- Il modello relazionale
- Relazioni e operazioni su relazioni
- Il linguaggio SQL
- Integrità delle basi di dati
 - Il concetto di transazione
 - Serializzazione delle transazioni

Le basi di dati

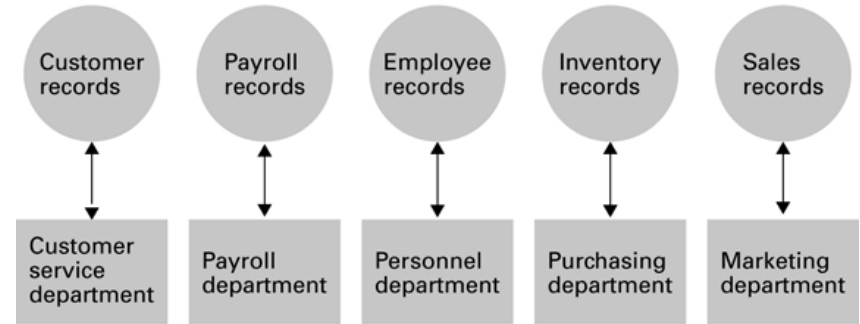
- Una **base di dati** è una collezione *multidimensionale* di dati, nel senso che i legami interni tra i suoi elementi permettono di accedere alle informazioni da una varietà di prospettive.
- In un file “piatto” possiamo memorizzare sequenzialmente una lista di opere ordinate per autore; in una base di dati con le stesse informazioni possiamo ottenere i lavori di un dato musicista, i compositori che hanno scritto un certo tipo di musica, ecc...
- In una Base di Dati (o **DB, DataBase**) le informazioni sono “strutturate” in modo da
 - garantire una buona “qualità” dei dati;
 - consentire “interrogazioni” efficienti

Organizzazione con uso di file e con uso di basi di dati

Sistema informativo orientato a file: ogni applicazione ha i suoi dati, con grande ridondanza

Sistema informativo orientato a base di dati: i dati sono centralizzati e integrati, e tutte le applicazioni li condividono

a. File-oriented information system

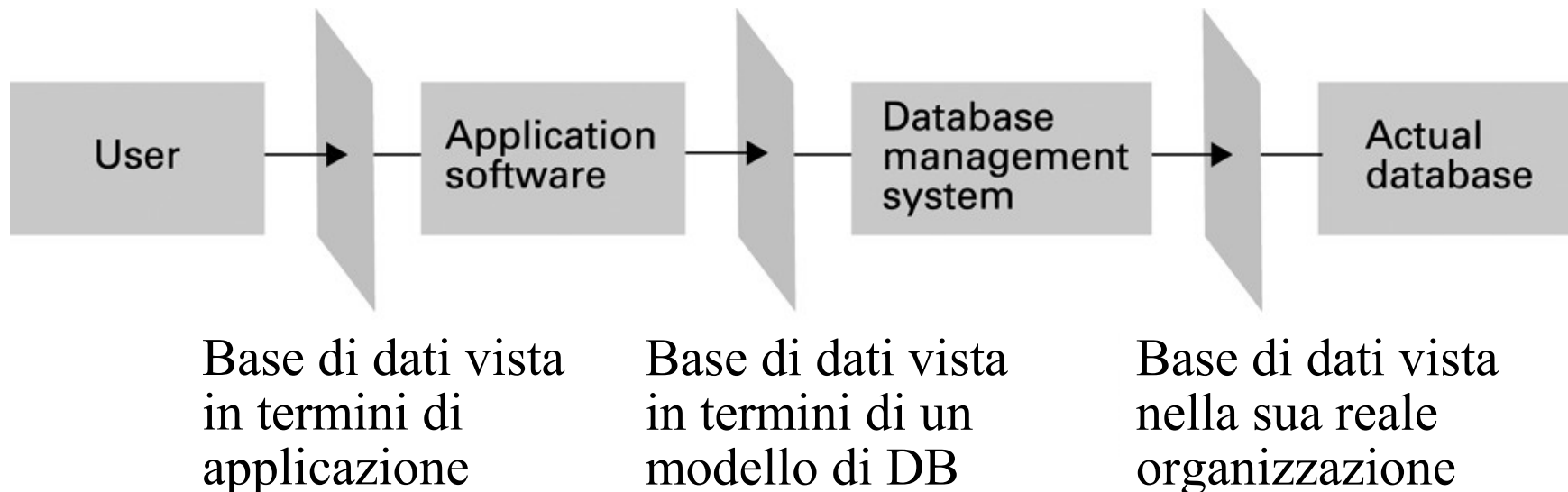


b. Database-oriented information system



I livelli concettuali dell'implementazione di una base di dati

- Per accedere ai dati di un DB si passano vari livelli
 - Il programmatore scrive programmi in un linguaggio di programmazione esteso con procedure per creare, modificare, cancellare dati, e fare interrogazioni
 - Queste procedure attivano funzionalità del Sistema di Gestione di Basi di Dati (DBMS) che le realizza operando sui dati reali



Schemi e sottoschemi

- **Schema:** Una descrizione della struttura di un intero database, usata dal Sistema di Gestione per mantenere la base di dati.
- **Sottoschema:** Una descrizione di una porzione della base di dati, relativa alle esigenze di una certa categoria di utenti. Consente di evitare che dati sensibili siano acceduti da persone non autorizzate.

Il sistema di gestione delle basi di dati

- **Database Management System (DBMS):** Un livello software che gestisce una base di dati, operando su di essa in risposta a richieste dalle applicazioni.
- **Database Distribuito:** Un database memorizzato su più computer:
 - Il DBMS nasconderà questo dettaglio implementativo agli utenti
- **Indipendenza dai dati:** La capacità di cambiare l'organizzazione di un database senza cambiare le applicazioni software che la usano.

Modelli concettuali di basi di dati

- **Un modello di dati:** è una vista concettuale della base di dati.
 - Modello di dati relazionale
 - Modello di dati orientato agli oggetti
 - Modello gerarchico

Il modello di dati relazionale

- **Relazione:** Una tabella rettangolare
 - **Attributi (Campi/Field):** Una colonna della tabella
 - **Tupla (Record):** Un riga nella tabella

Una relazione contenente informazioni su impiegati

Empl Id	Name	Address	SSN
25X15	Joe E. Baker	33 Nowhere St.	111223333
34Y70	Cheryl H. Clark	563 Downtown Ave.	999009999
23Y34	G. Jerry Smith	1555 Circle Dr.	111005555
•	•	•	•
•	•	•	•
•	•	•	•

Progettazione di relazioni

- Progettazione di uno schema
- Evitare di catturare più concetti con la stessa relazione
 - Può portare a ridondanza di dati
 - Cancellando una tupla si potrebbero cancellare anche delle informazioni necessarie ma non collegate
- **Decomposizione:** Dividere le colonne di una relazione in due o più relazioni, duplicando le colonne necessarie per mantenere la relazione
 - Decomposizione **senza perdita di dati:** una decomposizione “corretta” che non perde informazione

Un database su impiegati costituito da tre relazioni

EMPLOYEE relation

Empl Id	Name	Address	SSN
25X15	Joe E. Baker	33 Nowhere St.	111223333
34Y70	Cheryl H. Clark	563 Downtown Ave.	999009999
23Y34	G. Jerry Smith	1555 Circle Dr.	111005555

JOB relation

Job Id	JobTitle	Skill Code	Dept
S25X	Secretary	T5	Personnel
S26Z	Secretary	T6	Accounting
F5	Floor manager	FM3	Sales
.	.	.	.
.	.	.	.
.	.	.	.

ASSIGNMENT relation

Empl Id	Job Id	Start Date	Term Date
23Y34	S25X	3-1-1999	4-30-2006
34Y70	F5	10-1-2007	*
23Y34	S26Z	5-1-2006	*
.	.	.	.
.	.	.	.
.	.	.	.

Interrogazioni: come trovare i reparti dove l'impiegato 23Y34 ha lavorato

EMPLOYEE relation

Empl Id	Name	Address	SSN
25X15	Joe E. Baker	33 Nowhere St.	111223333
34Y70	Cheryl H. Clark	563 Downtown Ave.	999009999
23Y34	G. Jerry Smith	1555 Circle Dr.	111005555
.	.	.	.
.	.	.	.
.	.	.	.

JOB relation

Job Id	Job Title	Skill Code	Dept
S25X	Secretary	T5	Personnel
S26Z	Secretary	T6	Accounting
F5	Floor manager	FM3	Sales
.	.	.	.
.	.	.	.
.	.	.	.

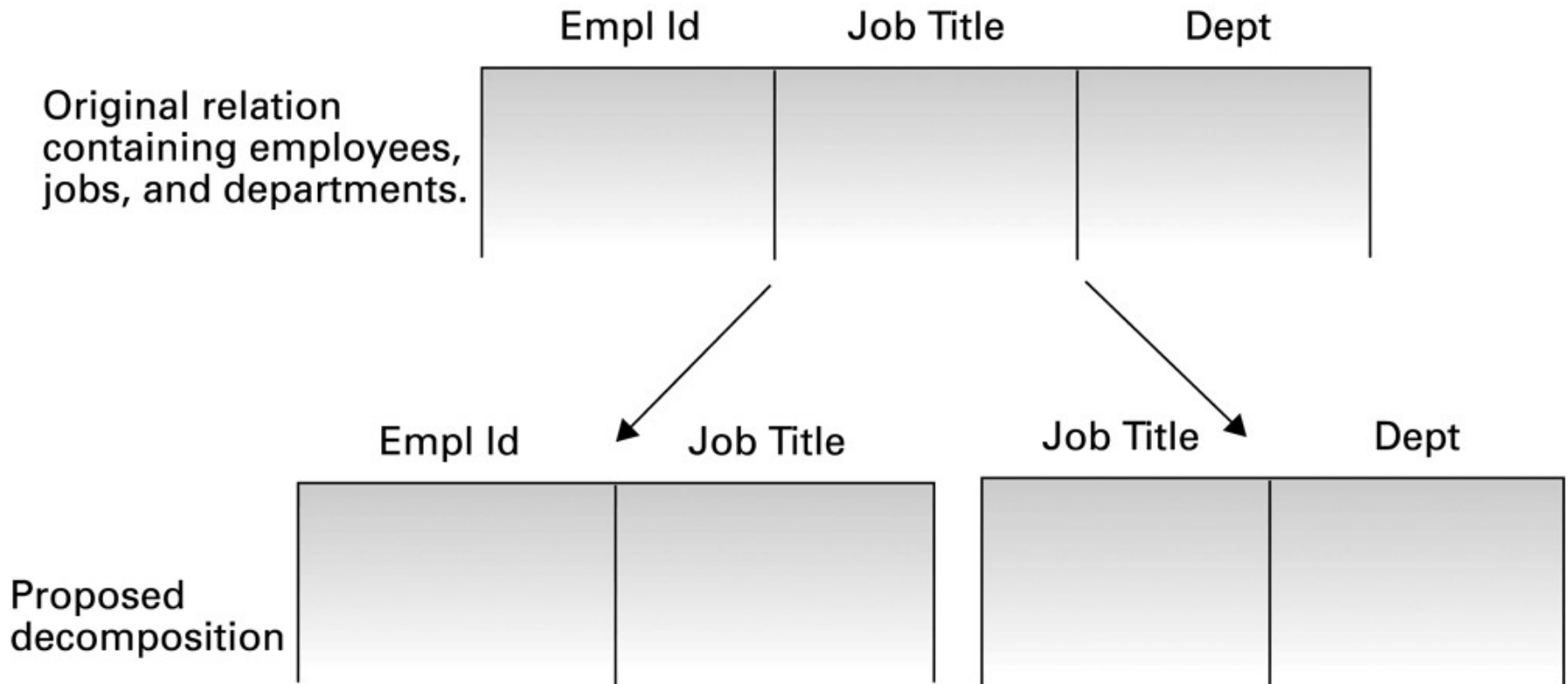
are contained in the personnel and accounting departments.

ASSIGNMENT relation

Empl Id	Job Id	Start Date	Term Date
23Y34	S25X	3-1-1999	4-30-2006
34Y70	F5	10-1-2007	*
23Y34	S26Z	5-1-2006	*
.	.	.	.
.	.	.	.
.	.	.	.

The jobs held by employee 23Y34

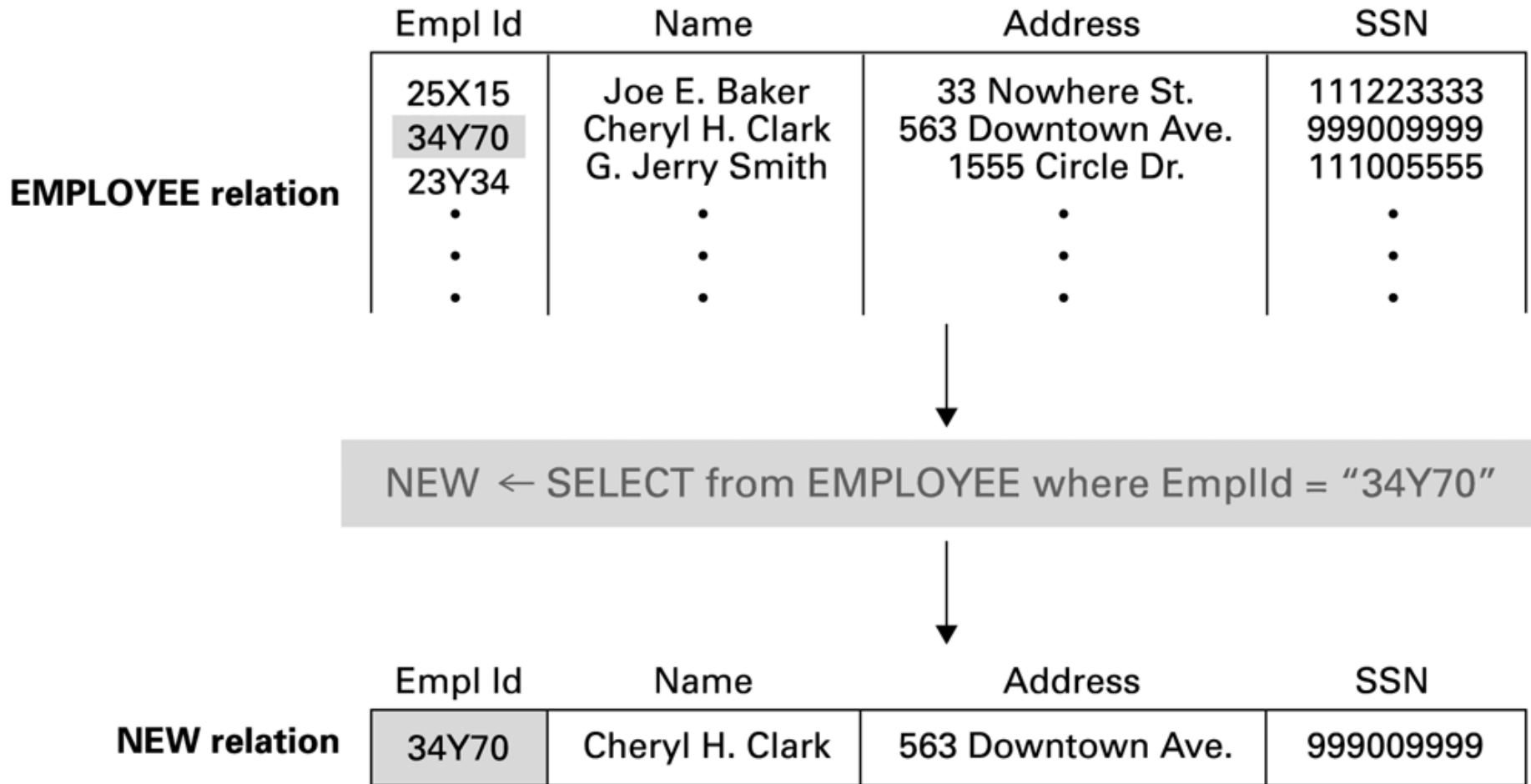
Una relazione e una possibile decomposizione



Operazioni su relazioni (calcolo relazionale)

- **Select:** Sceglie le tuple (righe della tabella)
- **Project:** Sceglie le colonne
- **Join:** Combina informazioni tra due o più relazioni.

L'operazione SELECT



L'operazione PROJECT

EMPLOYEE relation

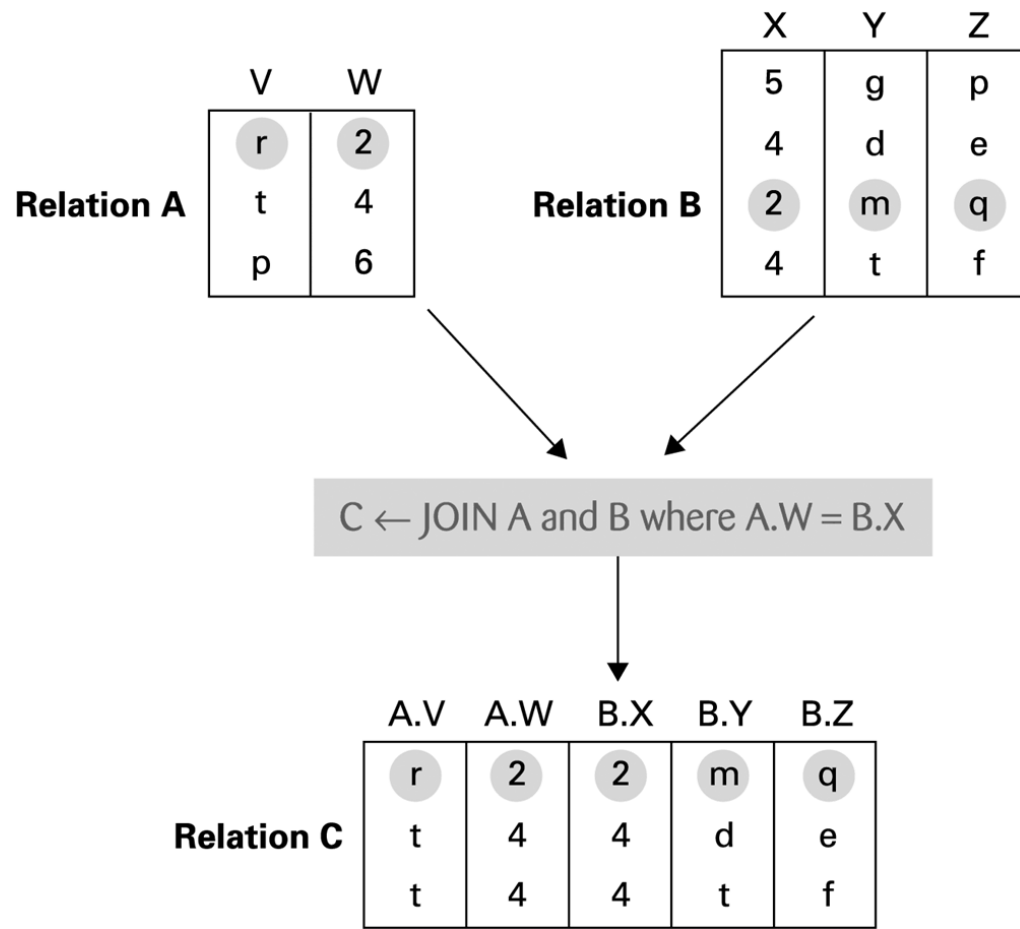
Empl Id	Name	Address	SSN
25X15	Joe E. Baker	33 Nowhere St.	111223333
24Y70	Cheryl H. Clark	563 Downtown Ave.	999009999
23Y34	G. Jerry Smith	1555 Circle Dr.	111005555
•	•	•	•
•	•	•	•
•	•	•	•

MAIL ← PROJECT Name, Address from EMPLOYEE

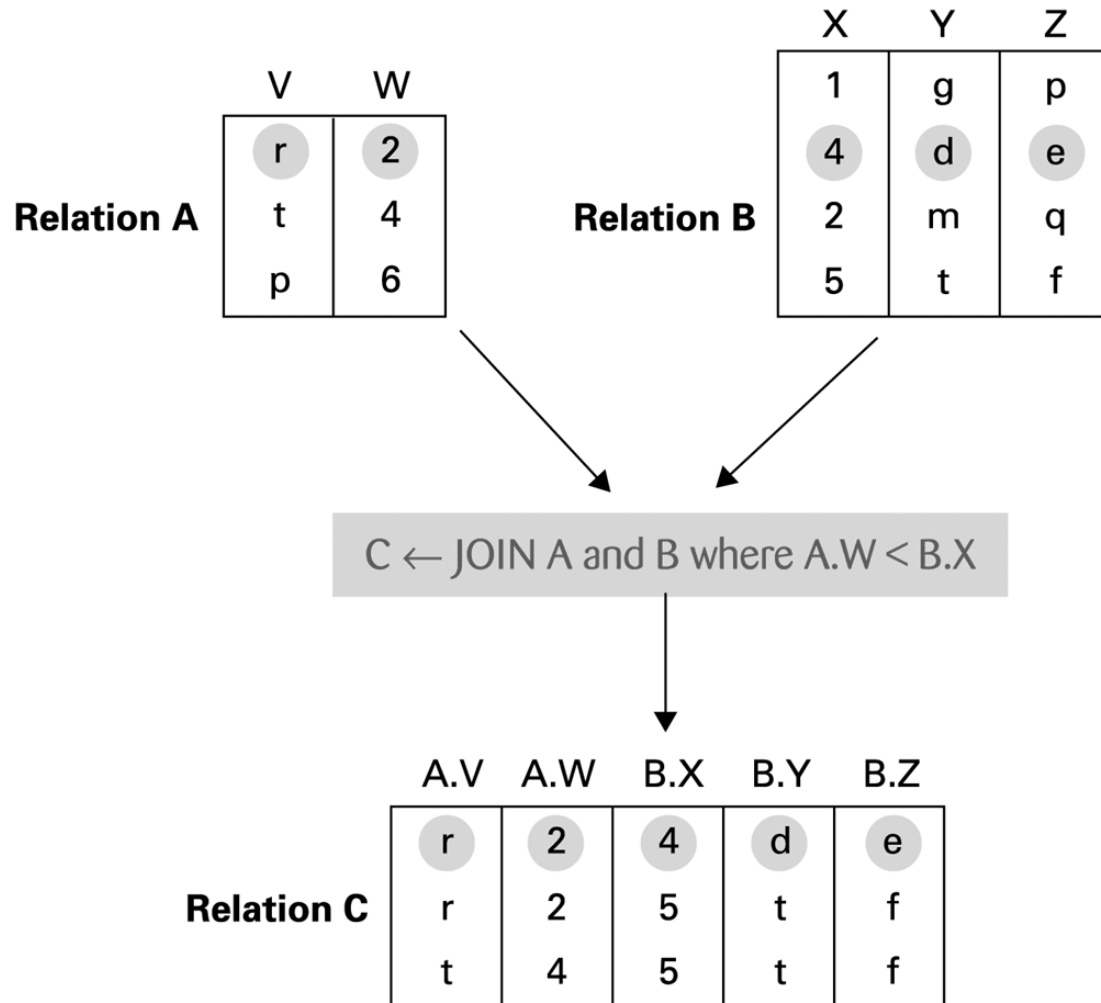
MAIL relation

Name	Address
Joe E. Baker	33 Nowhere St.
Cheryl H. Clark	563 Downtown Ave.
G. Jerry Smith	1555 Circle Dr.
•	•
•	•
•	•

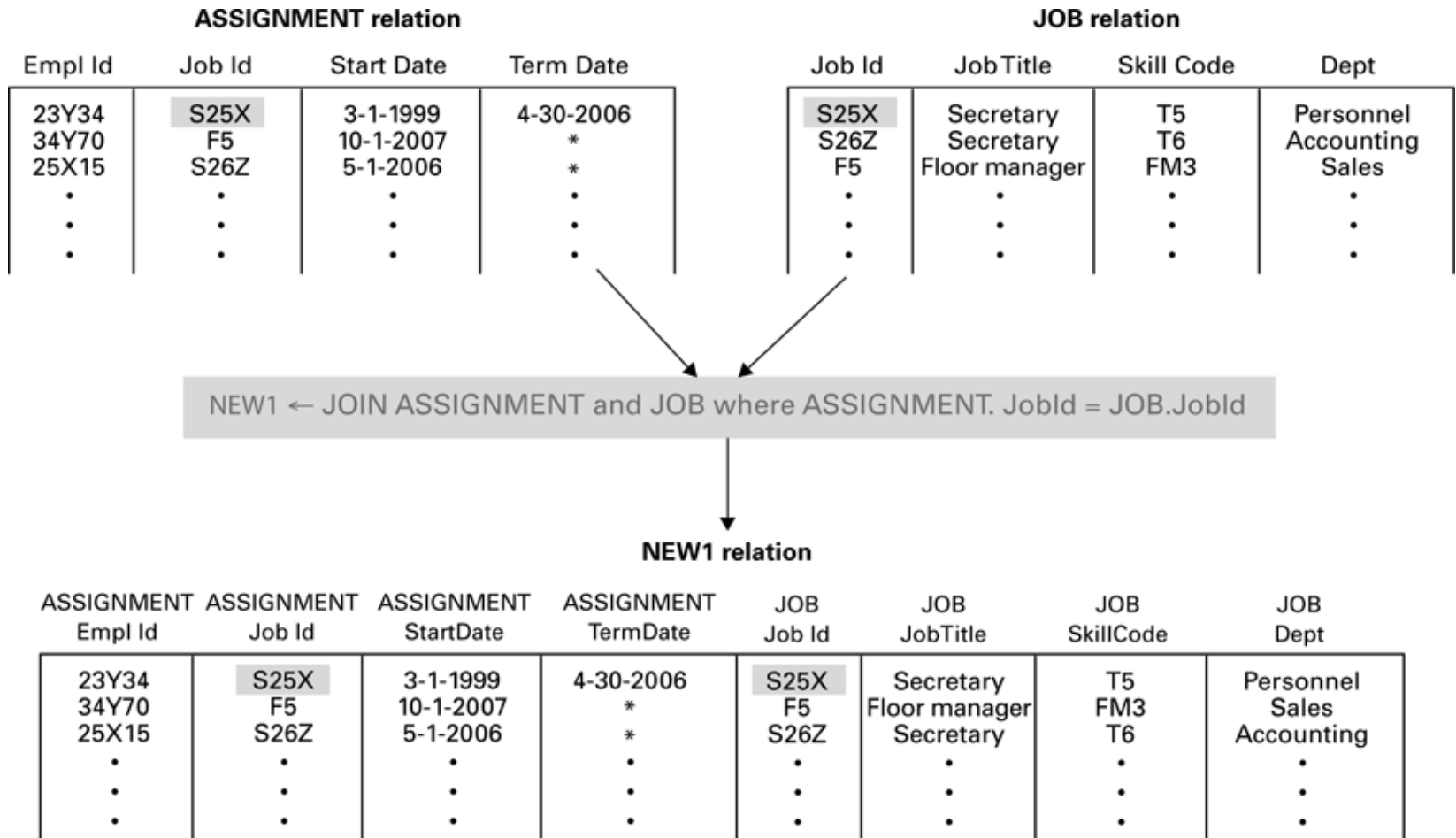
L'operazione JOIN



Un altro esempio di JOIN



Un'applicazione dell'operazione JOIN



Structured Query Language (SQL)

- Operazioni per manipolare tuple:
 - inserimento
 - aggiornamento
 - cancellazione
 - select: definizione di nuove relazioni ottenute come combinazione di quelle esistenti (con SELECT, PROJECT e JOIN)

Esempi SQL

- ```
select EmplId, Dept
from ASSIGNMENT, JOB
where ASSIGNMENT.JobId = JOB.JobId
 and ASSIGNMENT.TermData = "*"

```
- ```
insert into EMPLOYEE
values ('43212', 'Sue A. Burt',
      '33 Fair St.', '444661111')

```

Esempi SQL

- `delete from EMPLOYEE
where Name = 'G. Jerry Smith'`
- `update EMPLOYEE
set Address = '1812 Napoleon Ave.'
where Name = 'Joe E. Baker'`

Gestione dell'integrità dei database

- **Transazione:** Una sequenza di operazioni che devono essere eseguite tutte insieme
 - Esempio: trasferimento di danaro da un conto bancario ad un'altro.
- **Log (registro) delle transazioni:** Un registrazione non volatile di tutte le attività di ogni transazione, costruita prima che la transazione sia autorizzata ad eseguire
 - **Punto di Commit:** Il punto in cui la transazione è stata registrata nel log.
 - **Roll-back:** Il processo di annullare una transazione

Gestione dell'integrità dei database

- Problemi dovuti all'accesso concorrente:
 - Il problema della somma incorretta
 - Il problema dell'aggiornamento perso
- **Locking** = impedire ad altri di accedere ai dati che sono usati da una transazione
 - **Lock condiviso**: usato quando si leggono i dati
 - **Lock esclusivo**: usato quando si modificano i dati