

Basi di Dati

Barbara Catania

1 Aspetti Introduttivi

Organizzazione corso

- Aspetti introduttivi
- Modelli tradizionali dei dati: il modello relazionale
 - algebra e calcolo relazionale
 - il linguaggio SQL
- Progettazione di basi dati
 - metodologie di progettazione
 - il modello Entity-Relationship
 - elementi di teoria della normalizzazione
- Cenni all'architettura di un DBMS

2 Aspetti Introduttivi

Organizzazione corso

- Lezioni di laboratorio:
 - Microsoft SQL Server
- Esame:
 - compiti (scritto/laboratorio)
 - prova scritta/prova di laboratorio
 - orale obbligatorio solo in caso di sufficienza scarsa su determinati argomenti

3 Aspetti Introduttivi

**Basi di Dati e Sistemi Informativi:
Aspetti Introduttivi**

4 Aspetti Introduttivi

Organizzazione/Azienda

- Utilizza un insieme di **risorse** e **regole** per lo svolgimento coordinato delle attività al fine del perseguimento degli scopi
- le risorse di una azienda (o amministrazione):
 - persone
 - denaro
 - materiali
 - **informazioni**

5 Aspetti Introduttivi

Sistema informativo

- Componente (sottosistema) di ogni organizzazione che gestisce i processi informativi (cioè i processi che coinvolgono le informazioni)
- Acquisisce, elabora, conserva, produce le informazioni di interesse (cioè utilizzate per il perseguimento degli scopi dell'organizzazione stessa)
- ogni organizzazione ha un sistema informativo, eventualmente non esplicitato nella struttura
- il sistema informativo è di solito suddiviso in sottosistemi (in modo gerarchico o decentrato), più o meno fortemente integrati

6 Aspetti Introduttivi

Funzioni di un sistema informativo

- raccolta, acquisizione delle informazioni
- archiviazione, conservazione delle informazioni
- elaborazione delle informazioni
- distribuzione, scambio di informazioni

7

Aspetti Introduttivi

Sistemi informativi e automazione

- il concetto di “sistema informativo” è indipendente da qualsiasi automatizzazione: esistono organizzazioni la cui ragion d’essere è la gestione di informazioni (es: servizi anagrafici e banche) e che operano da secoli

8

Aspetti Introduttivi

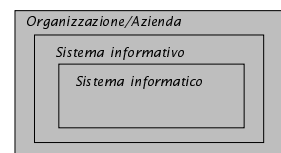
Sistema Informatico

- porzione automatizzata del sistema informativo
- il sistema informatico è la parte del sistema informativo che gestisce informazioni per mezzo della tecnologia informatica

9

Aspetti Introduttivi

Sistema Informatico



10

Aspetti Introduttivi

Sistema informativo e Sistema Informatico

- anche prima di essere automatizzati, molti sistemi informativi si sono evoluti verso una razionalizzazione e standardizzazione delle procedure e dell'organizzazione delle informazioni

11

Aspetti Introduttivi

Gestione delle informazioni

- Nelle attività umane, le informazioni vengono gestite (registrate e scambiate) in forme diverse, a seconda delle necessità e capacità:
 - idee informali
 - linguaggio naturale (scritto o parlato, formale o colloquiale, in una lingua o in un'altra)
 - disegni, grafici, schemi
 - numeri
 - codici (anche segreti)
- e su vari supporti, dalla memoria umana alla carta

12

Aspetti Introduttivi

Gestione delle informazioni

- Nelle attività standardizzate dei sistemi informativi complessi, sono state introdotte col tempo forme di organizzazione e codifica delle informazioni
- ad esempio, nei servizi anagrafici si è iniziato con registrazioni discorsive e sono state poi introdotte informazioni via via più precise:
 - nome e cognome
 - estremi anagrafici
 - codice fiscale

13

Aspetti Introductivi

Informazioni e dati

- Nei sistemi informatici (e non solo in essi) le informazioni vengono rappresentate attraverso i dati
 - **informazione**: tutto ciò che produce variazioni nel patrimonio conoscitivo di un soggetto detto *percettore dell'informazione*
 - **dato**: una registrazione della descrizione di una qualsiasi caratteristica della realtà su un supporto che ne garantisca la conservazione e, mediante un insieme di simboli, ne garantisca la comprensibilità e la reperibilità

14

Aspetti Introductivi

Informazioni e dati

- I dati sono fatti elementari, informazioni codificate, che hanno bisogno di essere interpretate per fornire conoscenza
- *Esempio*:
 - si consideri un dato rappresentato dal numero 3: tale dato non fornisce in effetti alcuna informazione
 - viceversa, dire che 3 è il numero degli anni di corso del corso di Laurea in Informatica fornisce una informazione

15

Aspetti Introductivi

Informazioni e dati

- Uno degli obiettivi fondamentali di un sistema di gestione dati è fornire un contesto interpretativo ai dati, in modo da consentire un accesso efficace alle informazioni da essi rappresentate

16

Aspetti Introductivi

Base di dati

(accezione generica)

- **collezione di dati, utilizzati per rappresentare le informazioni di interesse per una o più applicazioni di una organizzazione**

(accezione specifica)

- **collezione di dati gestita da un DBMS**

17

Aspetti Introductivi

Sistema di Gestione di Basi di Dati DataBase Management System — DBMS

Sistema (**prodotto software**) in grado di gestire **collezioni di dati** che siano:

- **Grandi**: di dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati
- **Persistenti**: con un periodo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano
- **Condivise**: utilizzate da applicazioni diverse

18

Aspetti Introductivi

DBMS (continua)

- Un DBMS deve garantire:
 - **Affidabilità**: resistenza a malfunzionamenti hardware e software
 - **Privatezza**: con una disciplina e un controllo degli accessi
- come ogni prodotto informatico, un DBMS deve essere **efficiente** (utilizzando al meglio le risorse di spazio e tempo del sistema) ed **efficace** (rendendo produttive le attività dei suoi utilizzatori)

19

Aspetti Introduttivi

Condivisione

- Ogni organizzazione (specie se grande) è divisa in settori o comunque svolge diverse attività
- a ciascun settore o attività corrisponde un (sotto-)sistema informativo (privato o porzione di un sistema più grande)
- possono esistere sovrapposizioni fra i dati di interesse dei vari settori
- una base di dati è una risorsa **integrata**, condivisa fra i vari settori

20

Aspetti Introduttivi

Condivisione (segue)

- L'integrazione e la condivisione permettono di ridurre la **ridondanza** (evitando ripetizioni) e, di conseguenza, le possibilità di incoerenza (o **inconsistenza**) fra i dati
- poiché la condivisione non è mai completa (o comunque non opportuna) i DBMS prevedono meccanismi per garantire **privatezza** ai dati (**autorizzazioni**)
- la condivisione richiede un opportuno coordinamento degli accessi: **controllo della concorrenza**

21

Aspetti Introduttivi

Efficienza

- Si misura (come in tutti i sistemi informatici) in termini di tempo di esecuzione (tempo di risposta) e spazio di memoria (principale e secondaria)
- i DBMS, a causa della varietà di funzioni, non sono necessariamente più efficienti dei file system
- l'efficienza è il risultato della qualità del DBMS e delle applicazioni che lo utilizzano

22

Aspetti Introduttivi

DBMS vs file system

- La gestione di insiemi di dati grandi e persistenti è possibile anche attraverso sistemi più semplici — gli ordinari **file system** dei sistemi operativi, che permettono di realizzare anche rudimentali forme di condivisione
- i DBMS estendono le funzionalità dei file system, fornendo più servizi ed in maniera integrata

23

Aspetti Introduttivi

DBMS vs file system

- Si supponga che una società voglia mantenere informazioni sui suoi impiegati e dipartimenti e si supponga che le applicazioni usino direttamente i servizi del file system per la memorizzazione e l'accesso ai dati
- in base a tale approccio, i dati relativi agli impiegati ed ai dipartimenti sono mantenuti in record memorizzati in file su memoria secondaria

24

Aspetti Introduttivi

DBMS vs file system

- Si supponga che esista un insieme di programmi applicativi tra cui:
 - un programma di modifica dello stipendio di un dato impiegato
 - un programma di modifica del dipartimento di un dato impiegato
 - un programma per l'inserzione e cancellazione di un impiegato
 - un programma che stampa la lista di tutti gli impiegati in ordine alfabetico

25

Aspetti Introduttivi

DBMS vs file system: problemi

- Ridondanze e inconsistenze nei dati:
 - uffici diversi potrebbero mantenere file diversi
 - i dati di uno stesso impiegato potrebbero essere replicati
 - a fronte di una modifica (es. cambia lo stipendio)
 - tutte le repliche devono essere aggiornate
 - problemi di inconsistenza se questo non avviene (lo stesso impiegato potrebbe essere associato a stipendi diversi)
 - in un DBMS si mantiene una sola rappresentazione, riducendo al massimo ridondanza e inconsistenza

26

Aspetti Introduttivi

DBMS vs file system: problemi

- Difficoltà nell'accesso ai dati:
 - ogni tipologia di ricerca corrisponde ad un nuovo programma applicativo
 - ogni nuova esigenza comporta quindi la necessità di scrivere un nuovo programma
 - i DBMS permettono in modo semplice di esprimere condizioni di ricerca sui dati, senza dovere scrivere complessi programmi applicativi

27

Aspetti Introduttivi

DBMS vs file system: problemi

- Problemi nell'accesso concorrente ai dati:
 - la gestione della concorrenza a livello di file system avviene a livello di file
 - un DBMS gestisce la concorrenza a livello di singolo dato acceduto
- Problemi di protezione dei dati:
 - il file system ci permette di proteggere un singolo file
 - in un DBMS possiamo proteggere in modo diverso varie porzioni di dati contenuti in un singolo file

28

Aspetti Introduttivi

DBMS vs file system: problemi

- Problemi di integrità dei dati:
 - se i dati devono soddisfare determinate condizioni (ad esempio, si vuole imporre che lo stipendio sia un numero positivo), ogni programma applicativo deve forzare questa condizione
 - in un DBMS, è possibile specificare una sola volta i vincoli di integrità, che verranno automaticamente verificati dal sistema

29

Aspetti Introduttivi

Obiettivi di un DBMS

- | | | |
|----------------------------------|---|--|
| Integrazione dei dati | → | Aumentare la possibilità di consultare i dati, ridurre ridondanza e inconsistenza |
| Indipendenza dei dati | → | Agevolare lo sviluppo di nuove applicazioni, facilitare la manutenzione delle applicazioni esistenti |
| Controllo centralizzato sui dati | → | Garantire qualità, privacy e sicurezza dei dati |

30

Aspetti Introduttivi

Obiettivi di un DBMS

Integrazione dei dati

- Il meccanismo fondamentale che rende possibile l'integrazione è una definizione logica e centralizzata dei dati detta **schema** della base di dati
- uno schema descrive tramite un formalismo ad alto livello, detto **modello dei dati**, il contenuto della base di dati

31

Aspetti Introduttivi

Obiettivi di un DBMS

Indipendenza dei dati

- obiettivo estremamente importante in quanto garantisce che:
 - modifiche apportate alla rappresentazione fisica dei dati (ad esempio uso di una struttura di accesso piuttosto di un'altra) non comportino modifiche alle applicazioni esistenti - **indipendenza fisica**
 - modifiche apportate alla rappresentazione logica dei dati non comportino modifiche alle applicazioni esistenti - **indipendenza logica**

32

Aspetti Introduttivi

Obiettivi di un DBMS

Controllo centralizzato sui dati

- Il terzo obiettivo consiste nell'introdurre un controllo centralizzato sui dati tramite il DBMS che fornisce un metodo comune e controllato di accesso alla base di dati

33

Aspetti Introduttivi

Servizi offerti da un DBMS

Servizio	Descrizione
Descrizione dei dati	Per specificare i dati da memorizzare
Manipolazione dei dati	Per: accedere ai dati, memorizzare nuovi dati, modificare/cancellare dati esistenti
Controllo di integrità	Per evitare di memorizzare dati non corretti
Strutture di memorizzazione	Per rappresentare in memoria i concetti di un modello dei dati

34

Aspetti Introduttivi

Servizi offerti da un DBMS

Servizio	Descrizione
Ottimizzazione di interrogazioni	Per determinare la strategia più efficiente per accedere ai dati
Privatizza dei dati	Per proteggere i dati da accessi non autorizzati
Ripristino del DB	Per evitare che errori e malfunzionamenti creino un DB inconsistente
Dizionario dei dati	Per determinare i dati disponibili nel DB

35

Aspetti Introduttivi

Modello dei dati

- Un modello dei dati è un insieme di strumenti concettuali, o *formalismo*, che consta di tre componenti fondamentali:
 - un insieme di strutture dati
 - una notazione per specificare i dati tramite le strutture dati del modello
 - un insieme di operazioni per manipolare i dati

36

Aspetti Introduttivi

Modello dei dati

- I modelli dei dati permettono di rappresentare la realtà di interesse per un certo insieme di applicazioni
- concetti rilevanti alla base di tale rappresentazione:
 - **Entità**: "oggetto" della realtà applicativa di interesse, esistente e distinguibile da tutti gli altri
 - **Attributo**: una proprietà significativa, ai fini della descrizione della realtà applicativa di interesse, di una data entità
 - Ogni entità è pertanto caratterizzata da uno o più attributi; un attributo assume uno o più valori, detti *valori dell'attributo*, da un insieme di valori possibili detto *dominio dell'attributo*

37

Aspetti Introduttivi

Modello dei dati

- **Insieme di entità**: raggruppa un insieme di "oggetti" con le stesse caratteristiche; in altre parole, raggruppa entità "simili" aventi gli stessi attributi, anche se non necessariamente gli stessi valori per tali attributi
- Un insieme di attributi i cui valori identificano univocamente una entità in un dato insieme di entità costituisce una **chiave** per l'insieme di entità
- **Associazione**: una corrispondenza tra gli elementi di due (o più) insiemi di entità

38

Aspetti Introduttivi

Modello dei dati: Esempio

- Si consideri la realtà applicativa relativa agli impiegati ed ai dipartimenti:
 - **entità**: Mario Rossi, il dipartimento 30
 - **insiemi di entità**: l'insieme di tutti gli impiegati
l'insieme di tutti i dipartimenti
 - **attributi**: nome di impiegato, stipendio, mansione, numero di dipartimento, nome di dipartimento
 - **associazione**: il fatto che Mario Rossi lavora nel dipartimento 30

39

Aspetti Introduttivi

Modello dei dati

- Qualsiasi modello dei dati deve rispondere a due domande fondamentali:
 - (a) come rappresentare le entità e i loro attributi
 - (b) come rappresentare le associazioni
- la maggioranza dei modelli usa strutture come il record, in cui ogni componente rappresenta un attributo
- i modelli differiscono notevolmente in questo aspetto; la rappresentazione può avvenire tramite:
 - strutture, valori, puntatori

40

Aspetti Introduttivi

Il modello relazionale

- E' basato su una singola struttura dati - la *relazione*
- una relazione può essere vista come una tabella con righe, dette *tuple*, e colonne contenenti dati di tipo specificato, come ad esempio interi e stringhe

41

Aspetti Introduttivi

Tabelle: rappresentazione di relazioni

CORSI	Corso	Docente	Aula
	Basi di dati	Rossi	DS3
	Sistemi	Neri	N3
	Reti	Bruni	N3
	Controlli	Bruni	G

AULE	Nome	Edificio	Piano
	DS1	Ex-OMI	Terra
	N3	Ex-OMI	Terra
	G	Pincherle	Primo

42

Aspetti Introduttivi

Schema e istanze

In ogni base di dati esistono:

- lo **schema**, sostanzialmente invariante nel tempo, che ne descrive la struttura (aspetto **intensionale**); nell'esempio, le intestazioni delle tabelle
- l'**istanza**, costituita dai valori attuali, che possono cambiare rapidamente (aspetto **estensionale**); nell'esempio, il "corpo" di ciascuna tabella

43

Aspetti Introduttivi

Schema e istanze

- Il primo passo nello sviluppo di una base di dati è la definizione dello schema della base di dati
- successivamente vengono immessi i dati veri e propri che devono conformarsi alla definizione data dallo schema
- l'insieme dei dati presente in dato momento in una base di dati è detto *insieme delle istanze* della base di dati

44

Aspetti Introduttivi

Due tipi (principali) di modelli

- **modelli logici**: utilizzati per l'organizzazione dei dati; ad essi fanno riferimento i programmi; sono indipendenti dalle strutture fisiche; esempi: relazionale, reticolare, gerarchico, a oggetti
- **modelli concettuali**: permettono di rappresentare i dati in modo indipendente da ogni sistema, cercando di descrivere i concetti del mondo reale; sono utilizzati nelle fasi preliminari di progettazione; il più noto è il modello **Entity-Relationship**

45

Aspetti Introduttivi

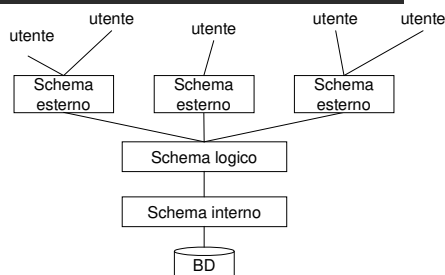
Modelli logici: Evoluzione

- Modelli di prima generazione (anni 60):
 - reticolare – CodasyI
 - gerarchico
- Modello relazionale (anni 70)
- Evoluzioni recenti:
 - modelli relazionali estesi (basi di dati attive)
 - modelli dei dati ad oggetti
 - modelli di basi di dati deduttive
 - modelli relazionali ad oggetti

46

Aspetti Introduttivi

Architettura ANSI/SPARC a tre livelli per DBMS



47

Aspetti Introduttivi

Architettura ANSI/SPARC: schemi

- schema logico**: descrizione dell'intera base di dati nel modello logico "principale" del DBMS
- schema esterno**: descrizione di parte della base di dati in un modello logico ("viste" parziali, derivate, anche in modelli diversi)
- schema fisico**: rappresentazione dello schema logico per mezzo di strutture fisiche di memorizzazione

48

Aspetti Introduttivi

Una vista

Corsi			Aule		
Corso	Docente	Aula	Nome	Edificio	Piano
Basi di dati	Rossi	DS3	DS1	Ex-OMI	Terra
Sistemi	Neri	N3	N3	Ex-OMI	Terra
Reti	Bruni	N3	G	Pincherle	Primo
Controlli	Bruni	G			

CorsiSedi			
Corso	Aula	Edificio	Piano
Sistemi	N3	Ex-OMI	Terra
Reti	N3	Ex-OMI	Terra
Controlli	G	Pincherle	Primo

49

Aspetti Introduttivi

Indipendenza dei dati

- Conseguenza della articolazione in livelli: l'accesso avviene solo tramite il livello esterno (che può coincidere con il livello logico)
- **indipendenza fisica**: il livello logico e quello esterno sono indipendenti da quello fisico; una relazione è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica (che può anche cambiare nel tempo senza che debbano essere modificate le forme di utilizzo)

50

Aspetti Introduttivi

Indipendenza dei dati

- **indipendenza logica**: il livello esterno è indipendente da quello logico
 - aggiunte o modifiche alle viste non richiedono modifiche al livello logico
 - modifiche allo schema logico che lascino inalterato lo schema esterno sono trasparenti

51

Aspetti Introduttivi

DBMS: linguaggi

- **Data Definition Language (DDL)**. Definisce:
 - lo schema logico e le viste del DB
 - vincoli di integrità
 - autorizzazioni per l'accesso ai dati
- **Data Manipulation Language (DML)**
 - utilizzato per l'interrogazione (query language) e l'aggiornamento del DB
- **Storage Definition Language (SDL)**
 - definisce lo schema fisico del DB

52

Aspetti Introduttivi

Data Definition Language

- La descrizione dei dati, cioè la definizione dello schema della base di dati, viene realizzata mediante il DDL
- un DDL in particolare concretizza il modello dei dati fornendo la notazione che permette di specificarne le strutture
- il DDL deve supportare la specifica del nome della base di dati, come pure di tutte le unità logiche elementari della base di dati (ad esempio relazioni e colonne nel caso del modello relazionale) e di eventuali vincoli di integrità semantica

53

Aspetti Introduttivi

Data Manipulation Language

- Una base di dati, viene acceduta dagli utenti e dalle applicazioni tramite il DML
- le operazioni supportate da questo linguaggio sono fondamentalmente quattro:
 - *Inserimento*: per l'immissione di nuovi dati
 - *Ricerca*: per il ritrovamento dei dati di interesse. Una operazione di ricerca è spesso detta *interrogazione* da cui il nome *linguaggio di interrogazione* del DML che permette di esprimere le operazioni di ricerca
 - *Cancellazione* - per l'eliminazione di dati obsoleti
 - *Aggiornamento* - per la modifica di dati

54

Aspetti Introduttivi

Storage Definition Language

- Le corrispondenze fra le strutture logiche dei dati, specificate nello schema della base di dati, e le strutture di memorizzazione devono essere opportunamente definite
- nella maggior parte dei DBMS attuali la definizione di tali corrispondenze è eseguita automaticamente dal DBMS stesso una volta che lo schema è definito
- tuttavia l'utente esperto può influenzare le scelte operate dal DBMS richiedendo ad esempio l'allocazione di strutture ausiliarie di accesso (come i B-alberi) per velocizzare determinati accessi ai dati
- tali richieste vengono effettuate tramite i comandi del linguaggio di definizione delle strutture di memorizzazione

55

Aspetti Introduttivi

SQL, un linguaggio interattivo

```
SELECT Corso, Aula, Piano
FROM Aule, Corsi
WHERE Nome = Aula
AND Piano="Terra";
```

Corso	Aula	Piano
Reti	N3	Terra
Sistemi	N3	Terra

56

Aspetti Introduttivi

SQL immerso in Pascal (o altro linguaggio ad alto livello)

```
write('nome della città?'); readln(citta);
EXEC SQL DECLARE P CURSOR FOR
  SELECT NOME, REDDITO
  FROM PERSONE
  WHERE CITTÀ = :citta ;
EXEC SQL OPEN P ;
EXEC SQL FETCH P INTO :nome, :reddito ;
while SQLCODE = 0 do begin
  write('nome della persona:', nome, 'aumento?');
  readln(aumento);
  EXEC SQL UPDATE PERSONE SET REDDITO = REDDITO + :aumento
  WHERE CURRENT OF P
EXEC SQL FETCH P INTO :nome, :reddito
end;
EXEC SQL CLOSE CURSOR P
```

57

Aspetti Introduttivi

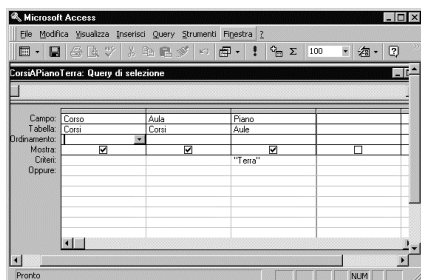
SQL immerso in linguaggio ad hoc (Oracle PL/SQL)

```
declare Stip number;
begin
  select Stipendio into Stip
  from Impiegato
  where Matricola = '575488'
  for update of Stipendio;
  if Stip > 30 then
    update Impiegato set Stipendio = Stipendio * 1.1 where Matricola = '575488';
  else
    update Impiegato set Stipendio = Stipendio * 1.15 where Matricola = '575488';
  end if;
  commit;
exception
  when no_data_found then
    insert into Errori
    values('Non esiste la matricola specificata',sysdate);
end;
```

58

Aspetti Introduttivi

Interazione non testuale (in Access)



59

Aspetti Introduttivi

Personaggi e interpreti

- progettisti e realizzatori di DBMS
- progettisti della base di dati e amministratori della base di dati (DBA)
- progettisti e programmatori di applicazioni
- utenti
 - utenti finali (terminalisti): eseguono applicazioni predefinite (**transazioni**)
 - utenti casuali: eseguono operazioni non previste a priori, usando linguaggi interattivi

60

Aspetti Introduttivi

Database administrator (DBA)

- Persona o gruppo di persone responsabile del controllo centralizzato e della gestione del sistema, delle prestazioni, dell'affidabilità, delle autorizzazioni
- le funzioni del DBA includono anche (parte di) quelle di progettazione

61

Aspetti Introduttivi

Transazioni

- Programmi che realizzano attività frequenti e predefinite, con poche eccezioni, previste a priori
 - versamento presso uno sportello bancario
 - emissione di certificato anagrafico
 - dichiarazione presso l'ufficio di stato civile
 - prenotazione aerea
- Le transazioni sono di solito realizzate con programmi in linguaggio ospite (tradizionale o ad hoc)
- **N.B.:** il termine **transazione** ha un'altra accezione, più specifica: sequenza indivisibile di operazioni (o vengono eseguite tutte o nessuna)

62

Aspetti Introduttivi