Basi di Dati - I

Turno M-Z Raimondo Schettini

AA. 2006/07

3/16/2007

Atzeni, Ceri, Paraboschi, Torlone Basi di dati McGraw-Hill, 1999-2002

Capitolo 1: Introduzione

AA 2006/2007

Sistemi informativi e sistemi

informatici

una premessa

- Scienza del trattamento razionale, specialmente

dell'informazione, considerata come supporto alla conoscenza umana e alla comunicazione

• Che cos'è l'informatica? Una definizione:

per mezzo di macchine automatiche,

Risorse

- le risorse di una azienda (o amministrazione):
 - persone
 - denaro
 - materiali
 - informazioni

Base di Dati: Insieme organizzato di dati utilizzati per il supporto allo svolgimento di attività (di un ente, azienda, ufficio, persona)

3/16/2007

· L'informatica ha due anime:

(Academie Francaise).

- tecnologica: i calcolatori elettronici e i sistemi che li utilizzano;
- metodologica: i metodi per la soluzione di problemi e la gestione delle informazioni.

3/16/2007

Sistema informativo

- Componente (sottosistema) di una organizzazione che gestisce le informazioni di interesse (cioè utilizzate per il perseguimento degli scopi dell'organizzazione)
- · funzioni del sistema informativo
 - acquisizione/memorizzazione
 - aggiornamento
 - interrogazione
 - elaborazione

3/16/2007

Esempio: Il sistema informativo dell'Università UNIVERSITA' DEGLI STUDI CHISSADOVE Corso di Studi in Ingegneria Informatica UNIVERSITA' DEGLI STUDI ROMA TRE Orario di ricevimento dei docenti Corso di Studi in Ingegneria Informatica Lunedi' 12-13 Giovedi' 11-13 Martedi' 16-18 Luigi NERI Piero ROSS • acquisizione/memorizzazione · aggiornamento interrogazione • elaborazione

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 (p. es. servizi anagrafici e banche) e che operano da secoli.
- Anche prima di essere automatizzati, molti sistemi informativi si sono evoluti verso una razionalizzazione e standardizzazione delle procedure e dell'organizzazione delle informazioni

3/16/2007

Sistema Informatico

 porzione automatizzata del sistema informativo: la parte del sistema informativo che gestisce informazioni con tecnologia informatica



.

Sistema Informatico Gestisce un sistema informativo in modo automatizzato

- •Garantisce che i dati siano conservati in modo permanente sui dispositivi di memorizzazione
- •Permette un rapido aggiornamento dei dati per riflettere rapidamente le loro variazioni
- •Rende i dati accessibili alle interrogazioni degli utenti
- •Può essere distribuito sul territorio

3/16/2007

Gestione delle informazioni

- Nelle attività umane, le informazioni vengono gestite (registrate e scambiate) in forme diverse:
 - idee informali
 - linguaggio naturale (scritto o parlato, formale o colloquiale, in una lingua o in un'altra)
 - disegni, grafici, schemi
 - numeri e codici
- e su vari supporti
 - memoria umana, carta, dispositivi elettronici

3/16/2007 10

Gestione delle informazioni

- Nelle attività standardizzate dei sistemi informativi complessi, sono state introdotte col tempo forme di organizzazione e codifica delle informazioni via via più precise (e in un certo senso artificiali).
- Ad esempio, nei servizi anagrafici si è iniziato con registrazioni discorsive e poi
 - nome e cognome
 - estremi anagrafici
 - codice fiscale

3/16/2007 11

Informazioni e dati

- Nei sistemi informatici (e non solo), le informazioni vengono rappresentate in modo essenziale, spartano: attraverso i dati
- Dal Vocabolario della lingua italiana (1987)
 informazione: notizia, dato o elemento che
 consente di avere conoscenza più o meno esatta
 di fatti, situazioni, modi di essere.
 - dato: ciò che è immediatamente presente alla conoscenza, prima di ogni elaborazione; (in informatica) elementi di informazione costituiti da simboli che debbono essere elaborati.

Dati e informazioni

• I dati hanno bisogno di essere interpretati

Esempio

'Mario' '275' su un foglio di carta sono due dati. Se il foglio di carta viene fornito in risposta alla domanda "A chi mi devo rivolgere per il problema X; qual è il suo numero di telefono?", allora i dati possono essere interpretati per fornire informazione e arricchire la conoscenza.

• La rappresentazione precisa di forme più ricche di informazione e conoscenza è difficile

3/16/2007 13

Basi di dati

- · L'elemento chiave è il dato
 - La rappresentazione precisa di forme più ricche di informazione e conoscenza è difficile
 - I dati costituiscono spesso una risorsa strategica, perché più stabili nel tempo di altre componenti (processi, tecnologie, ruoli umani)
 - I dati rimangono gli stessi nella migrazione da un sistema al successivo

3/16/2007 14

Base di dati

(accezione generica, metodologica)

 Insieme organizzato di dati utilizzati per il supporto allo svolgimento delle attività di un ente (azienda, ufficio, persona)

(accezione specifica, metodologica e tecnologica)

• insieme di dati gestito da un DBMS



3/16/2007

Base di dati

Ancora una definizione:

insieme di archivi in cui ogni dato e' rappresentato logicamente una sola volta e puo' essere utilizzato da un insieme di applicazioni da diversi utenti secondo opportuni criteri di riservatezza.





16

Gestione di base di dati

SW specifico scritti in un liguaggio di programmazione generico (C, C++, COBOL, Java)

DBMS

- Access
- DB2
- Oracle
- Informix
- Sybase
- SQLServer
- Ingres

3/16/2007

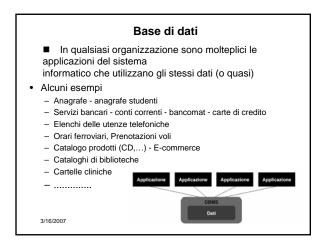
17

Sistema di gestione di basi di dati **DataBase Management System — DBMS**

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

- 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)

garantendo **affidabilità** (resistenza a malfunzionamenti hardware e software) e **privatezza** (con una disciplina e un controllo degli accessi). Come ogni prodotto informatico, un DBMS deve essere <u>efficiente</u> (utilizzando al meglio le risorse di spazio e tempo del sistema) ed <u>efficace</u> (rendendo produttive le attività dei suoi utilizzatori).



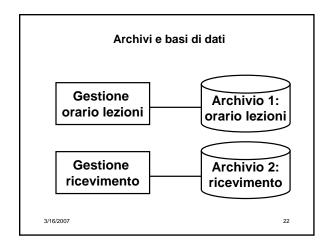
Condivisione

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

3/16/2007

20







Le basi di dati sono condivise Una base di dati e' una risorsa integrata, condivisa fra le varie applicazioni conseguenze Attivita' diverse su dati in parte condivisi: meccanismi di autorizzazione l'utente A è autorizzato a leggere tutti i dati e a modificare quelli sul ricevimento l'utente B è autorizzato a leggere i X e a modificare Y Attivita' multi-utente su dati condivisi: controllo della concorrenza L'effetto di transazioni concorrenti deve essere coerente

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.

Efficacia

- •Difficile da quantificare.
- •Una applicazione e' efficace se risponde agli obiettivi per cui e' stata progettata

3/16/2007 25

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
- I file system prevedono forme rudimentali di condivisione: "tutto o niente"
- I DBMS estendono le funzionalità dei file system, fornendo più servizi ed in maniera integrata.
- Nei programmi tradizionali che accedono a file, ogni programma contiene una descrizione della struttura del file stesso, con i conseguenti rischi di incoerenza fra le descrizioni (ripetute in ciascun programma) e i file stessi
- Nei DBMS, esiste una porzione della base di dati (il catalogo o dizionario) che contiene una descrizione centralizzata dei dati, che può essere utilizzata dai vari programmi

/2007 26

Descrizioni dei dati nei DBMS

- Descrizioni e rappresentazioni dei dati a livelli diversi
 - permettono l'indipendenza dei dati dalla rappresentazione fisica:
 - i programmi fanno riferimento alla struttura a livello più alto, e le rappresentazioni sottostanti possono essere modificate senza necessità di modifica dei programmi
 - Precisiamo attraverso il concetto di
 - · modello dei dati

3/16/2007 27

Modello dei dati

- insieme di costrutti utilizzati per organizzare i dati di interesse e descriverne la dinamica
- componente fondamentale: meccanismi di strutturazione (o costruttori di tipo)
- come nei linguaggi di programmazione esistono meccanismi che permettono di definire nuovi tipi, così ogni modello dei dati prevede alcuni costruttori
- ad esempio, il modello relazionale prevede il costruttore relazione, che permette di definire insiemi di record omogenei

3/16/2007 28



Organizzazione dei dati in una base di dati

Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00

3/16/2007

30

Lo schema della base di dati

Orario

	Insegnamento	Docente	Aula	Ora
ļ	Analisi matem. I	Luigi Neri	N1	8:00
	Basi di dati	Piero Rossi	N2	9:45
	Chimica	Nicola Mori	N1	9:45
	Fisica I	Mario Bruni	N1	11:45
	Fisica II	Mario Bruni	N3	9:45
	Sistemi inform.	Piero Rossi	N3	8:00

L'istanza della base di dati

3/16/2007

31

Schemi e istanze

- In ogni base di dati esistono:
 - lo schema, sostanzialmente invariante nel tempo, che ne descrive la struttura, il significato (aspetto intensionale).
 - nell'esempio, le intestazioni delle tabelle
 - l'istanza, i valori attuali, che possono cambiare anche molto rapidamente (aspetto estensionale)
 - nell'esempio, il "corpo" di ciascuna tabella

3/16/2007 32

Due tipi (principali) di modelli

- modelli logici: utilizzati nei DBMS esistenti per l'organizzazione dei dati
 - utilizzati dai programmi
 - indipendenti dalle strutture fisiche esempi: **relazionale**, reticolare, gerarchico, a oggetti
- modelli concettuali: permettono di rappresentare i dati in modo indipendente da ogni sistema
 - cercano di descrivere i concetti del mondo reale
 - sono utilizzati nelle fasi preliminari di progettazione

il più noto è il modello Entity-Relationship

3/16/2007

33

35

Architettura (semplificata) di un DBMS utente Schema logico Schema interno

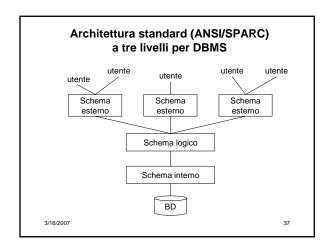
Architettura semplificata di un DBMS: schemi

- schema logico: descrizione della base di dati nel modello logico (ad esempio, la struttura della tabella)
- schema interno (o fisico): rappresentazione dello schema logico per mezzo di strutture memorizzazione (file)

3/16/2007

Indipendenza dei dati

- il livello logico è indipendente da quello fisico:
 - una tabella è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica (che può anche cambiare nel tempo)
- In questo corso vedremo solo il livello logico e non quello fisico



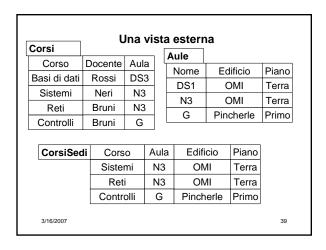
Architettura ANSI/SPARC: schemi

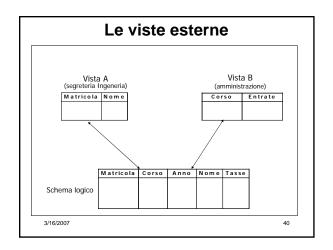
schema logico: descrizione dell'intera base di dati nel modello logico "principale" del DBMS

schema interno (fisico): rappresentazione dello schema logico per mezzo di strutture fisiche di memorizzazione

schema esterno: descrizione di parte della base di dati in un modello logico ("viste" parziali, derivate, anche in modelli diversi)

3/16/2007 38





Indipendenza dei dati

- · conseguenza della articolazione in livelli
- l'accesso avviene solo tramite il livello esterno (che può coincidere con il livello logico)
- due forme:
 - indipendenza fisica
 - indipendenza logica

3/16/2007 41

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
 - la realizzazione fisica può cambiare senza che debbano essere modificati i programmi

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

3/16/2007

Linguaggi per basi di dati

· Linguaggi per la definizione dei dati Data Definition Languages - DDL definizione degli schemi logici, fisici e delle autorizzazioni di

· Linguaggi di manipolazione dei dati Data Manipulation Languages - DML interrogazione e aggiornamento delle basi di dati

Alcuni linguaggi come SQL (Structured Query Language) presentano integrate funzioni di entrambe le categorie

3/16/2007 44

Linguaggi per basi di dati

43

45

47

- Un altro contributo all'efficacia: disponibilità di vari linguaggi e interfacce diverse
 - ⇒ linguaggi testuali interattivi (SQL)
 - ⇒ comandi (come quelli del linguaggio interattivo) immersi in un linguaggio ospite (Pascal, C, Cobol, etc.)
 - un linguaggio ad hoc, con anche altre funzionalità (p.es. per grafici o stampe strutturate), anche con l'ausilio di strumenti di sviluppo (p. es. per la gestione di maschere)

3/16/2007

SQL come DDL (Data Definition Language)

• Definizione di una tabella (relazione)

CREATE TABLE orario (insegnamento CHAR(20), docente CHAR(20), aula CHAR(4), CHAR(5)) ora insegnamento docente aula

orario

3/16/2007

SQL come **DML** (Data Manipulation Language)

Interrogazione

SELECT Insegnamento, Docente FROM Orario WHERE insegnamento = 'Basi di dati'

> Produce una tabella (relazione) che soddisfa il criterio (clausola WHERE)

insegnamento doce Basi di dati Schettini

3/16/2007

SQL, un linguaggio interattivo Corsi Docente Aula Corso Basi di dati Rossi DS3 N3 Sistemi Neri N3 Reti Bruni Controlli Bruni G

aggio intorattivo					
	Aule				
	Nome	Edificio	Piano		
	DS1	OMI	Terra		
	N3	OMI	Terra		
	G	Pincherle	Primo		

• "Trovare i corsi tenuti in aule a piano terra"

SELECT Corso, Aula, Piano FROM Aule, Corsi WHERE Nome = Aula

AND Piano="Terra"

3/16/2007

Corso Aula Piano Sistemi N3 Terra Reti N3 Terra

48



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 (<u>transazioni</u>)
 - utenti casuali: eseguono operazioni non previste a priori, usando linguaggi interattivi

3/16/2007 50

Database administrator (DBA)

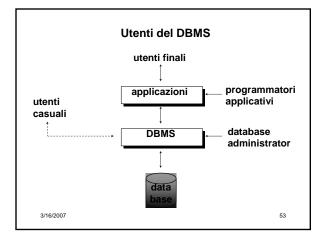
- Persona o gruppo di persone responsabile del controllo centralizzato e della gestione del sistema, delle prestazioni, dell'affidabilità, delle autorizzazioni
- Le funzioni del DBA includono quelle di progettazione, anche se in progetti complessi ci possono essere distinzioni

3/16/2007 51

Transazioni

- Programmi che realizzano attività frequenti e predefinite, con poche eccezioni, previste a priori.
- Esempi:
 - versamento presso uno presso 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).

3/16/2007 52



Riassunto

- •Un DB è una collezione di dati (ad es. rappresentati in forma di tabelle) che rappresentano le informazioni di interesse per un'organizzazione.
- •Progettare un DB significa trovare un "buon modo" per strutturare i dati, che rispetti sia i requisiti propri del Sistema Informativo dell'organizzazione sia i vincoli che sui dati vengono imposti.
- •Un DBMS è un sistema software che gestisce DB
- •Per eseguire operazioni su un DB è necessario scrivere istruzioni in un linguaggio (tipicamente SQL) supportato dal DBMS.
- •Gli operatori di cui dispone un DBMS sono vari, e variano da un sistema all'altro. Ogni DBMS ha comunque a disposizione un insieme di operatori che permettono di realizzare quelli propri dell'algebra relazionale.

Vantaggi e svantaggi dei DBMS

Pro

- gestione centralizzata dei dati come risorsa comune (a disposizione di tutta l'organizzazione - con opportune forme di controllo)
- disponibilità di servizi integrati
- riduzione di ridondanze e inconsistenze (qualita' dei dati)
- indipendenza dei dati (favorisce lo sviluppo e la manutenzione delle applicazioni)

Contro

- · costo dei prodotti e della transizione verso di essi
- non scorporabilità delle funzionalità (con riduzione di efficienza)

3/16/2007

Fine della presentazione

3/16/2007 56

Esercizi di autovalutazione

3/16/2007

57

55

· Quali delle seguenti affermazioni sono vere?

- l'indipendenza dei dati permette di scrivere programmi senza conoscere le strutture fisiche dei dati
- l'indipendenza dei dati permette di modificare le strutture fisiche dei dati senza dover modificare i programmi che accedono alla base di dati
- l'indipendenza dei dati permette di scrivere programmi conoscendo solo lo schema concettuale della BD
- l'indipendenza dei dati permette di formulare interrogazioni senza conoscere le strutture fisiche

3/16/2007 58

· Quali delle seguenti affermazioni sono vere?

- il fatto che le basi di dati siano condivise favorisce l'efficienza dei programmi che le utilizzano
- il fatto che le basi di dati siano condivise permette di ridurre ridondanze e inconsistenze
- il fatto che le basi di dati siano persistenti ne garantisce l'affidabilità
- il fatto che le basi di dati siano persistenti favorisce l'efficienza dei programmi
- il fatto che le basi di dati siano condivise rende necessaria la gestione della privatezza e delle autorizzazioni

3/16/2007 59

Quali delle seguenti affermazioni sono vere?

- la distinzione fra DDL e DML corrisponde alla distinzione fra schema e istanza
- le istruzioni DML permettono di interrogare la base di dati ma non di modificarla
- le istruzioni DDL permettono di specificare la struttura della base di dati ma non di modificarla
- non esistono linguaggi che includono sia istruzioni DDL sia istruzioni DML
- SQL include istruzioni DML e DDL
- le istruzioni DML permettono di interrogare la base di dati e di modificarla

- Quali delle seguenti affermazioni sono vere?
 - gli utenti casuali utilizzano transazioni predefinite
 - i terminalisti utilizzano transazioni predefinite
 - gli utenti casuali progettano la base di dati
 - i progettisti del DBMS realizzano le transazioni che saranno utilizzate dai terminalisti
 - i progettisti della base di dati realizzano il DBMS
 - i progettisti delle applicazioni utilizzano la base di dati come progettata dal progettista del DBMS
 - i progettisti delle applicazioni utilizzano la BD come progettata dal progettista della BD

3/16/2007 61

- Illustrare, in modo sintetico ma chiaro, supponendo di rivolgersi ad un non esperto,le caratteristiche fondamentali delle basi di dati e il ruolo che esse giocano nei sistemi informativi.
- Discutere brevemente (meno di mezza pagina) la seguente affermazione: "i dati sono una risorsa per una organizzazione, e come tali vanno considerati anche separatamente dalle applicazioni che li utilizzano."
- Illustrare brevemente (non più di mezza pagina) il concetto di indipendenza dei dati.