

Corso di Informatica

Modulo T1

A1 – Archivi di dati

M. Malatesta A1-Archivi di dati-26

1
15/09/2012

Prerequisiti

- Dati semplici e strutturati
- Strutture dati concrete e astratte
- Struttura e operazioni sui file
- Cenni sulle memorie secondarie
- File system

M. Malatesta A1-Archivi di dati-26

2
15/09/2012

Introduzione

Una necessità sempre crescente è quella di memorizzare grandi masse di informazioni. Sorgono alcune questioni.

Come rendere efficienti:

- **reperimento**
 - **conservazione**
- delle informazioni?**

In questa Unità ci proponiamo di introdurre il concetto di **archivio** che consente di dare una esauriente risposta ai quesiti accennati.

M. Malatesta A1-Archivi di dati-26

3
15/09/2012

Informazioni generali

In questa Unità si introduce il concetto di **archivio**, che rappresenta una evoluzione del concetto noto di file, e si vedranno le tecniche per organizzare i dati in memoria di massa, al fine di agevolare la loro:

- **memorizzazione** (*conservazione delle informazioni nel tempo e nello spazio*)
- **ricerca** (*reperimento delle informazioni in base a dati criteri*).

M. Malatesta A1-Archivi di dati-26

4
15/09/2012

Archivio come struttura logica

Intuitivamente, con il termine di **archivio** indichiamo un insieme di **informazioni organizzate su memoria di massa** che:

- l'utente vede come un'unica entità
- sono indipendenti dal dispositivo e dal supporto
- costituisce una **struttura dati astratta** per l'utente finale
- l'utente può registrare e/o rileggere nel tempo.

Ad esempio.

ISBN	Autore	Editore	Titolo	Dispon.	Prezzo
00V1	Vegetti Mario	Il Saggiatore	Il coltello e lo stilo	Si	12.50
00A8	Arnheim Rudolf	Feltrinelli	Film come arte	No	12.00
10A1	Rich Elaine	Mc Graw-Hill	A.I.	Si	11.50
00M2	Marx Karl	U.T.E.T	Il Capitale	Si	25.00

M. Malatesta A1-Archivi di dati-26

5
15/09/2012

Record e attributi

Nella progettazione di un archivio è necessario definire quali sono gli **attributi** (o **campi**) che rappresentano i dati in oggetto. Ad ogni insieme di valori per gli attributi diamo il nome di **record** o **registrazioni**.

ISBN	Autore	Editore	Titolo	Dispon.	Prezzo
00V1	Vegetti Mario	Il Saggiatore	Il coltello e lo stilo	Si	12.50
00A8	Arnheim Rudolf	Film come arte	Feltrinelli	No	12.00
10A1	Rich Elaine	Mc Graw-Hill	A.I.	Si	11.50
00M2	Marx Karl	U.T.E.T	Il Capitale	Si	25.00

M. Malatesta A1-Archivi di dati-26

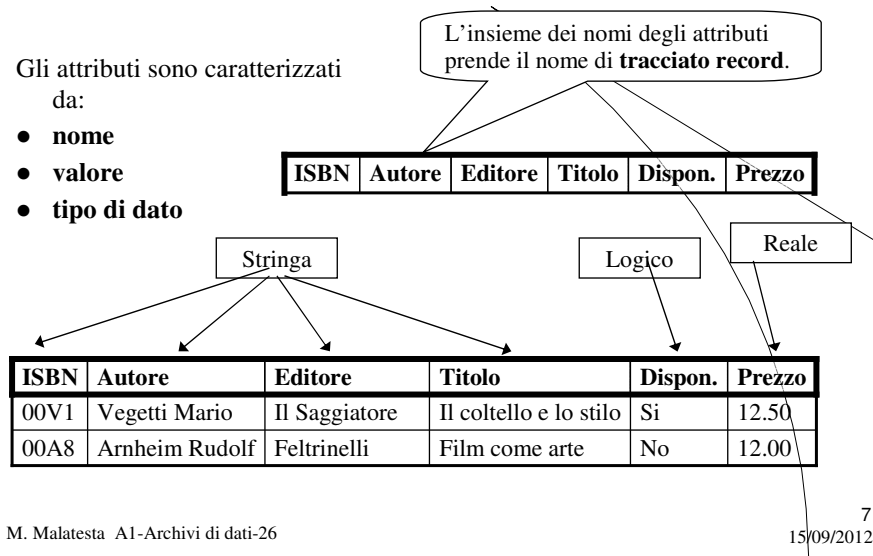
Esempio di struttura ad archivio

6
15/09/2012

Record e attributi

Gli attributi sono caratterizzati da:

- **nome**
- **valore**
- **tipo di dato**



Archivio come struttura fisica

Per memorizzare su memoria di massa un archivio, ci serviamo della struttura file. Un archivio può essere memorizzato:

- **su un singolo file**, se le informazioni sono “leggere”
- **su un gruppo di file**, in caso di informazioni più articolate.

Nel seguente esempio è mostrata la struttura logica di un archivio distribuito su 3 file.

Archivio come struttura fisica

CodEdit	Editore
45	Il Saggiatore
50	Feltrinelli
51	Mc Graw-Hill
53	U.T.E.T

CodAut	Autore
101	Vegetti Mario
201	Arnheim Rudolf
220	Rich Elaine
300	Marx Karl

ISBN	Autore	Editore	Titolo	Dispon.	Prezzo
00V1	101	45	Il coltello e lo stilo	Si	12.50
00A8	201	50	Film come arte	No	12.00
10A1	220	51	A.I.	Si	11.50
00M2	300	53	Il Capitale	Si	25.00

M. Malatesta A1-Archivi di dati-26

9
15/09/2012

Archivio come struttura fisica

Ad esempio, per l'archivio dell'esempio precedente, i gruppi di informazioni omogenee (ad esempio Autore, Editore) è opportuno siano registrati su file separati, poiché essendo presenti più volte:

- *utilizzano in modo inefficiente* lo spazio su disco (ripetizione)
- *possono creare malfunzionamenti* (ad es. variazione del nome dell'editore)

In pratica, un archivio viene memorizzato su memoria secondaria utilizzando uno o più *file* che ne costituiscono la **struttura fisica**.

Pertanto il **file** (o i **file**) rappresentano la visione (**concreta**) dell'archivio da parte del programmatore (**struttura fisica**), mentre l'**archivio** è la visione (**astratta**) da parte dell'**utente**.

M. Malatesta A1-Archivi di dati-26

10
15/09/2012

Record logico e fisico

La struttura che viene assegnata ai dati al momento della progettazione, e che abbiamo chiamato *tracciato record*, prende più precisamente il nome di **record logico**.

ISBN	Cognome	Nome	Editore	Titolo	Dispon.	Prezzo

In memoria di massa, ciascun record logico viene disposto fisicamente sul supporto dal *file system* sotto forma di **record fisico**.

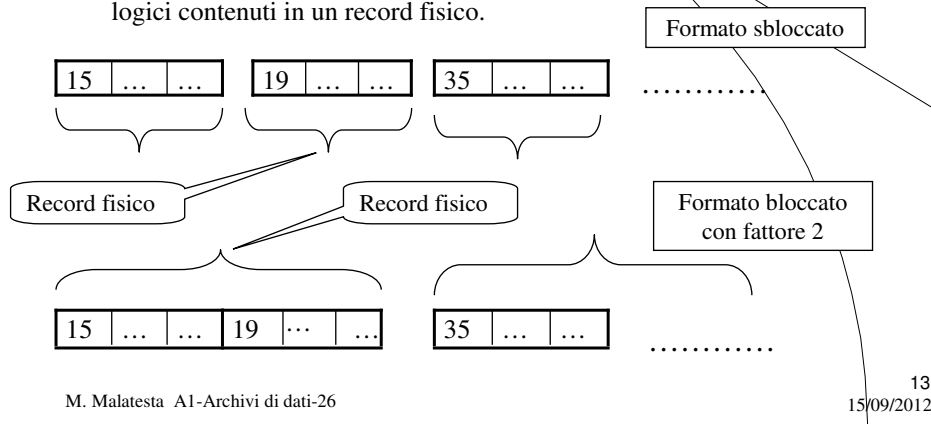
Record logico e fisico

Il *file system*:

- organizza i record logici in **record fisici** o **blocchi** che rappresentano l'unità minima di memorizzazione su disco o nastro;
- consente di **leggere** blocchi (*dal file alla RAM*) o **scrivere** blocchi (*dalla RAM sul file*);
- **dispone i blocchi secondo strategie** che ottimizzano i tempi di memorizzazione e di reperimento;
- **dimensiona il blocco** in base alle caratteristiche del supporto di memoria secondaria (nei sistemi comuni vanno da 32 a 4096 byte, il valore tipico è 512 byte);
- è in grado di **ricostruire l'informazione** iniziale, poiché mantiene i collegamenti fisici tra i blocchi contenenti i record logici appartenenti allo stesso file.

Fattore di bloccaggio

É possibile prevedere che in un record fisico vi siano più record logici; indichiamo con il termine **fattore di bloccaggio** il numero di record logici contenuti in un record fisico.



M. Malatesta A1-Archivi di dati-26

Buffer

I dati presenti in memoria di massa, per essere elaborati devono essere portati in RAM.

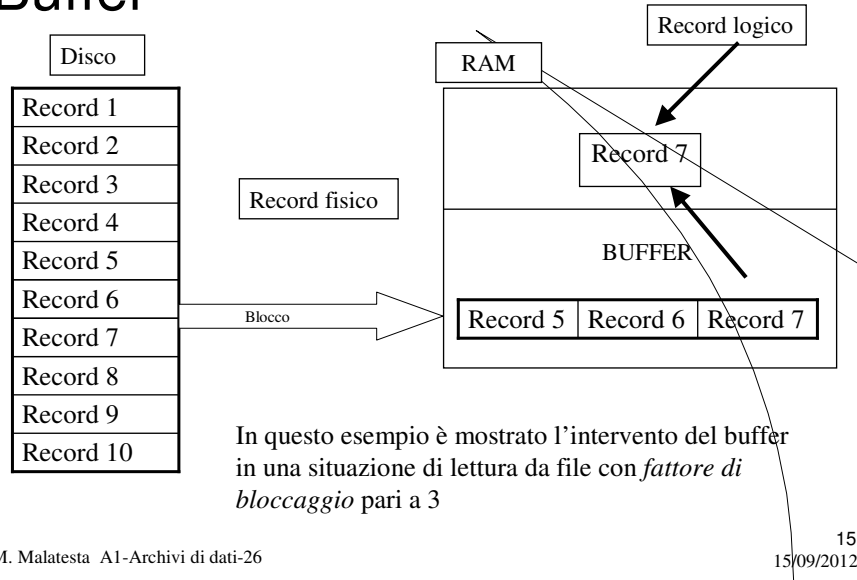
Poiché i tempi di accesso alla memoria secondaria per operazioni di lettura/scrittura sono molto maggiori rispetto a quelli di elaborazione (CPU), si cerca di minimizzare il numero di accessi al disco.

Le operazioni di lettura/scrittura che operano sui **record fisici** si appoggiano ad un'area in RAM detta **buffer** in cui transita il record fisico da leggere o scrivere.

M. Malatesta A1-Archivi di dati-26

14
15/09/2012

Buffer



M. Malatesta A1-Archivi di dati-26

Definizione di archivio

Un archivio è quindi un insieme di dati:

- **strutturato**, contenente cioè informazioni strutturate
- **memorizzato** permanentemente su memoria di massa, eventualmente mediante più di un file;
- **organizzato in modo razionale** per facilitare la loro gestione (ad es. mediante codici);
- **astratto** sulla quale si può agire con **operazioni logiche** di *creazione, ricerca e aggiornamento* (eliminazione, inserimento, modifica, ordinamento, ecc).

ATTIVITA': elencare gli insiemi di informazioni (file) necessarie per la gestione di:

- una scuola;
- un'agenzia viaggi
- un'attività commerciale

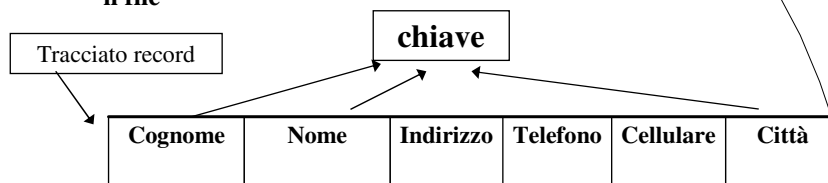
M. Malatesta A1-Archivi di dati-26

16
15/09/2012

Il concetto di chiave

Uno dei modi per organizzare razionalmente i dati in un archivio è quello di assegnare ai record dei file che lo compongono un codice detto **chiave**.

La chiave è un sottoinsieme degli attributi del record che costituisce il file



M. Malatesta A1-Archivi di dati-26

17
15/09/2012

Il concetto di chiave

Ad esempio:

- il codice fiscale in un archivio anagrafico;
- l'ISBN in un archivio di libri.(v. esempio precedente);
- il numero di conto corrente bancario;
- cognome, nome e città come nell'esempio precedente.

M. Malatesta A1-Archivi di dati-26

18
15/09/2012

Il concetto di chiave

Se la chiave *individua un solo record in modo univoco*, si dice **chiave primaria**.

- Ad es. il codice fiscale, l'ISBN, il numero di matricola di uno studente, il codice di un articolo in un magazzino

Se la chiave *individua un gruppo di record*, si dice chiave **secondaria**.

- Ad es. il cognome di una persona, la città di residenza, o il nome di una casa editrice.

ATTIVITA': disegnare il tracciato record evidenziando la chiave primaria, per ciascuno dei file necessari per la gestione di:

- una scuola;
- un'agenzia viaggi
- un'attività commerciale

La scelta della chiave primaria

La scelta della chiave primaria determina l'efficienza di qualunque operazione sull'archivio. Possiamo avere:

- **Chiave implicita**, in cui la chiave è una parte del tracciato record.
 - Ad esempio, il numero di telefono degli utenti di una compagnia telefonica
- **Chiave mediante tabella**, in cui si aggiunge al tracciato record un campo.
 - Ad esempio un **codice**, che individua il record
- **Chiave mediante regole logiche**, in cui la chiave è ottenuta da parti delle informazioni presenti nel record (codice parlante).
 - Ad esempio il codice fiscale.

Argomenti

- Archivio come struttura logica
- Record e attributi
- Archivio come struttura fisica
- Record logico e fisico
- Fattore di bloccaggio
- Buffer
- Definizione di archivio
- Il concetto di chiave
- La scelta della chiave primaria

M. Malatesta A1-Archivi di dati-26

21
15/09/2012

Altre fonti di informazione

- A.Lorenzi-D.Rossi, Le basi di dati e il linguaggio SQL – ed. ATLAS
- G.Callegarin – Corso di Informatica Generale 3 – CEDAM
- Cesarini, Pippolini, Soda – Informatica 3 – ed. Cremonese
- R.Dameri, G.Foresti, F.Percivale – Informatica. Vol 2 – ed. Jackson

M. Malatesta A1-Archivi di dati-26

22
15/09/2012