

# Corso di Informatica

## Modulo T1

### A2 – Organizzazione

M. Malatesta A2-Organizzazione-20

1  
28/09/2017

## Prerequisiti

- Strutture dati fondamentali
- Cenni sulle memorie secondarie
- Struttura file

M. Malatesta A2-Organizzazione-20

2  
28/09/2017

# Introduzione

Le tecnologie costruttive delle memorie secondarie sono evolute, partendo dai nastri fino a giungere agli attuali veloci hard disk e ai DVD che hanno grandissima diffusione.

Ci troviamo davanti ai seguenti quesiti:

**Come sono evolute, in corrispondenza, le tecnologie informatiche per consentire un'efficace sfruttamento dei supporti e realizzare elaborazioni veloci ed efficienti dei dati su memoria di massa?**

In questa Unità ci proponiamo di introdurre i concetti di **organizzazione** e di **accesso** agli archivi e di presentare alcune scelte tecnologiche, che l'analista dovrà di volta in volta valutare attentamente.

# Informazioni generali

In questa Unità vediamo come i dati possono essere disposti negli archivi in memoria di massa, anche seconda del tipo di supporto disponibile, per cui parleremo di **organizzazione degli archivi**.

Inoltre, in base all'organizzazione scelta, vedremo (nella successiva Unità) in quali modi possiamo reperire le informazioni registrate negli archivi, tecniche che vanno sotto il nome di **accesso agli archivi**.

# Organizzazione ed accesso

Nello studio degli archivi è importante considerare:

- **l'organizzazione degli archivi**, ossia il modo in cui le informazioni sono memorizzate;
- **le tecniche di accesso agli archivi**, ossia il modo in cui le informazioni possono essere reperite.

Nel seguito indichiamo con “archivio” l'insieme di uno o più file.

# Organizzazione

Con la frase **organizzazione degli archivi** si intende come i dati sono disposti nella memoria di massa e quali sono le operazioni con cui possiamo elaborarli.

Possiamo distinguere innanzitutto:

- **organizzazione fisica**
  - come i record fisici sono disposti sul supporto fisico
  - caratteristiche del tipo di supporto di memorizzazione scelto
  - visione dell'archivio da parte del progettista
- **organizzazione logica**
  - come i vari record logici sono disposti all'interno dell'archivio
  - visione dell'archivio da parte dell'utente

# Organizzazione fisica

L'**organizzazione fisica** dipende dalle caratteristiche del supporto.

Abbiamo due tipi di supporti:

- **supporti ad organizzazione sequenziale** – Tipicamente le **unità a nastro magnetico**, ma anche le unità a disco prevedono questo tipo di organizzazione. In questo caso organizzazione fisica e logica coincidono
- **supporti ad organizzazione casuale** – Tipicamente le **unità a disco** hanno questa caratteristica, ma possono svolgere anche le funzioni di supporto ad accesso sequenziale. In questo caso organizzazione fisica e logica possono essere molto diverse.

# Organizzazione fisica

## - supporti ad organizzazione sequenziale

Nei **nastri magnetici** l'organizzazione dei record può essere solo sequenziale.

- I record possono essere memorizzati e reperiti in ordine consecutivo ed è possibile fare inserimenti solo alla fine.
- Organizzazione fisica e logica coincidono
- Tra un record fisico e il successivo è presente uno *spazio vuoto, non registrabile* detto **gap**, che ha il compito di consentire la sincronizzazione delle operazioni di I/O. Poiché il nastro, per poter essere letto/scritto correttamente necessita di avanzare ad una *velocità costante* i gap servono a compensare le accelerazioni e le decelerazioni del supporto.

# Organizzazione fisica

## - supporti ad organizzazione sequenziale

Le caratteristiche fisiche di una unità a nastri sono:

- **Capacità (C)** – Il massimo numero di Gbyte memorizzabili sul nastro
- **Densità di registrazione (D)** – Il numero di byte per pollice (*byte/inch*) del nastro
- **Tempo di accesso ( $T_a$ )** – Tempo di attesa tra il record richiesto ed il suo trasferimento in memoria (a volte anche minuti)
- **Lunghezza (L)** – Lunghezza del supporto espressa in piedi (*feet*) (1 piede=12 pollici)
- **Velocità di avanzamento (V)** – Velocità di trascinamento del supporto (in *inch/sec*)
- **Lunghezza del gap ( $L_g$ )** – in pollici (*inch*)

# Organizzazione fisica

## - supporti ad organizzazione sequenziale

Le unità a nastro magnetico non hanno più grande diffusione, come era molti anni fa; sono spesso usate come unità di backup e di archiviazione.

**ATTIVITA'**: dato un file di 10.000 record, di 20 byte ciascuno, memorizzati a blocchi di 400 byte, calcolare il fattore di blocco e l'occupazione complessiva del file.

$$F_b = 400 / 20 = 20$$

$$N = 10.000 \times 20 = 200.000 \text{ byte} = 200 \text{ Kb}$$

# Organizzazione fisica

## - supporti ad organizzazione sequenziale

I nastri:

- sono supporti molto *economici*
- hanno una *grande capacità*
- presentano *alti tempi di elaborazione*.

**ATTIVITA'**: supponendo che il file precedente venga scandito alla velocità di 50 inch/sec, che la densità di registrazione sia 50 byte/inch e con gap lunghi 0.01 inch, calcolare il tempo medio per scandire l'intero file

$$T_m = 200.000 \text{ byte} / ( 50 \text{ inch/sec} \times 50 \text{ byte/inch} + 10.000/400 \times 0.01 \times 50 \text{ byte/inch} \times 50 \text{ inch/sec} ) = 64 \text{ sec}$$

M. Malatesta A2-Organizzazione-20

11  
28/09/2017

# Organizzazione fisica

## - supporti ad organizzazione casuale

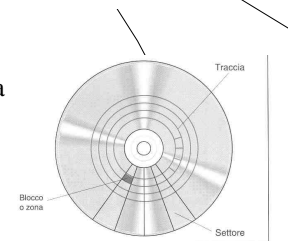
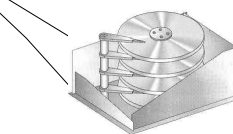
I **dischi magnetici fissi** sono costituiti da una *pila di dischi*, ciascuno formato da un *supporto* (plastico o metallico) ricoperto di materiale magnetico e da *testine* che permettono la *scrittura* e la *lettura* di bit su un dato disco, memorizzandoli sotto forma di **stati di polarizzazione** (positiva e negativa)

Ogni disco ha due **facce** ed è diviso in cerchi concentrici numerati, detti **tracce**. Ogni faccia del disco è suddivisa in spicchi detti **settori** anch'essi numerati. L'insieme *immaginario* delle tracce aventi lo stesso indirizzo si dice **cilindro**.

L'indirizzo di un settore è dato dalla combinazione di 3 indirizzi ordinati:

(*cilindro, faccia, settore*)

M. Malatesta A2-Organizzazione-20



12  
28/09/2017

# Organizzazione fisica

## - supporti ad organizzazione casuale

Il disco ruota ad una velocità fissa attorno al suo asse. Durante il reperimento di un dato, la testina di lettura/scrittura *seleziona elettronicamente il cilindro e la faccia desiderati* (con tempi dell'ordine dei msec), mediante un **accesso diretto**; successivamente, occorre attendere che il settore desiderato si porti sotto di essa (**accesso sequenziale**). In questo caso si parla di **accesso semidiretto**

Nei dischi l'organizzazione fisica e quella logica sono indipendenti e l'organizzazione dei record può essere sequenziale o in ordine sparso. Le operazioni di inserimento, aggiornamento e ricerca possono essere fatte in modo molto agevole.

Anche in questo caso è presente uno spazio vuoto, non registrabile detto **gap**, che ha il compito di separare i vari settori fisici delle tracce.

M. Malatesta A2-Organizzazione-20

13  
28/09/2017

# Organizzazione fisica

## - supporti ad organizzazione casuale

Le caratteristiche fisiche di una unità a disco sono:

- **Capacità (C)** – Il massimo numero di byte memorizzabili sul disco
- **Tempo di accesso ( $T_a$ )** – Tempo di attesa tra il record richiesto ed il suo trasferimento in memoria (sui millisecondi o meno)
- **Velocità di rotazione (V)** – in giri/minuto
- **Dimensione del disco** – in pollici

Altri parametri:

- numero di cilindri
- numero di tracce per cilindro
- numero di byte per traccia

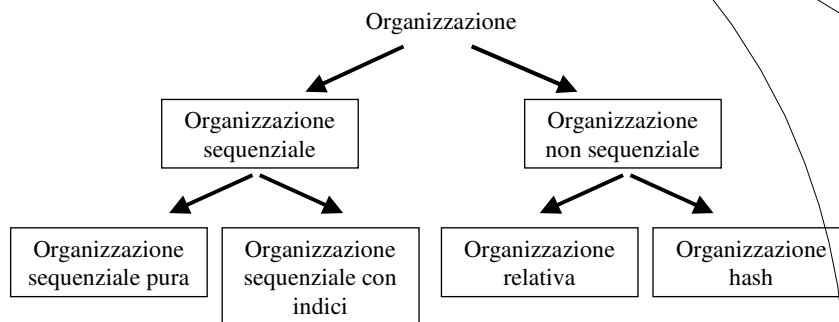


M. Malatesta A2-Organizzazione-20

14  
28/09/2017

# Organizzazione logica

L'**organizzazione logica** indica la disposizione dei record logici nell'archivio e le possibili operazioni per la loro gestione.



M. Malatesta A2-Organizzazione-20

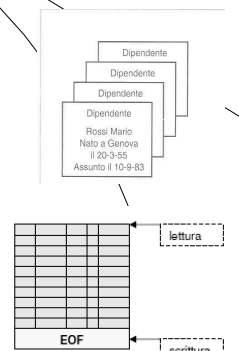
15  
28/09/2017

## Organizzazione logica - organizzazione sequenziale

L'**organizzazione sequenziale**:

- prevede che i record siano disposti ed acceduti in sequenza.
- è possibile con unità a nastro e a disco
- è da preferirsi quando si prevede uno stesso tipo di elaborazione su tutti i record
- è di solito abbastanza lenta
- consente **operazioni globali** (su gruppi di record)
- consente aggiunte **soltanto in coda** al file

Ad esempio, volendo gestire l'*operazione globale* di **mailing**, (invio di lettere personalizzate ai clienti) si può pensare di organizzare sequenzialmente l'archivio dei dati dei clienti.



M. Malatesta A2-Organizzazione-20

16  
28/09/2017



# Organizzazione logica

## - organizzazione sequenziale con indici

L'organizzazione sequenziale con indici

- prevede che i record siano individuati da una **chiave alfanumerica**, che viene associata automaticamente alla posizione a cui accedere per reperire il dato.
- è possibile solo nelle unità a disco;
- è di uso più facile per l'utente, che per ricercare un record può usare una chiave alfanumerica, anziché un numero;
- consente sia **operazioni locali** che **operazioni globali**.

M. Malatesta A2-Organizzazione-20

17  
28/09/2017

# Organizzazione logica

## - organizzazione sequenziale con indici

L'organizzazione ad indici prevede una implementazione fisica con un **file di indici** (*index file*) ed un **file di lavoro** (*work file*).

Il file di indici è formato da record in cui si hanno 2 campi:

- la chiave;
- l'indirizzo logico (posizione) del record associato nel file di lavoro.

Ad esempio, la chiave immessa dall'utente viene individuata tramite ricerca nell'index file. L'indirizzo associato alla posizione trovata consente accesso diretto nel work file all'informazione cercata

A20	5
A25	1
A26	2
B12	4
B13	3

File indici

1	A25	.....
2	A26	.....
3	B13	.....
4	B12	.....
5	A20	.....

File primario

Index file

Work file

M. Malatesta A2-Organizzazione-20

18  
28/09/2017

# Organizzazione logica

## - organizzazione sequenziale con indici

Nel file di lavoro sono presenti i dati che costituiscono le informazioni da elaborare; normalmente, questo file è gestito con accesso diretto mediante la posizione logica.

Il file ad indici è un file che in genere viene mantenuto ordinato in base alla chiave, in modo da poter utilizzare la ricerca binaria.

Per file non troppo voluminosi, l'index file può essere caricato in memoria centrale, per migliorare le prestazioni.

# Organizzazione logica

## - organizzazione non sequenziale (relativa)

L'**organizzazione relativa** (*random*)

- prevede che i record siano disposti in sequenza logica, ciascuno caratterizzato da un numero intero che ne rappresenta la posizione;
- è possibile solo nelle unità a disco;
- richiede che l'utente, per effettuare la ricerca, conosca la posizione del record da reperire;
- è piuttosto veloce;
- consente **operazioni locali** (su un record alla volta) e **globali**.

Ad esempio, nella gestione dei conti correnti, ciascun record può essere individuato proprio dal numero univoco di conto.

1				
2				
3				
N				

# Organizzazione logica

## - organizzazione non sequenziale (relativa)

In sostanza i record sono disposti come fossero in un array, con il vantaggio che i dati sono memorizzati su memoria di massa.

Quando si inserisce un record in una data posizione (assegnata dall'utente):

- se è disponibile la posizione, il record viene memorizzato;
- se la posizione non esiste, vengono creati tanti record vuoti fino a crearla, dopodichè il record viene inserito in essa.

# Organizzazione logica

## - organizzazione non sequenziale (hash)

L'**organizzazione hash**:

- consiste nell'applicare una formula  $f$  alla chiave alfanumerica  $K$  di un record per generarne l'indirizzo  $Ind$  all'interno di un file con organizzazione relativa;

$$Ind_K = f(K)$$

- non prevede che i record siano disposti in sequenza logica, poiché le posizioni sono distribuite in base alla scelta della formula;
- è possibile solo nelle unità a disco;
- richiede che l'utente, per effettuare la ricerca, conosca la chiave  $K$ ;
- è piuttosto veloce;
- consente **operazioni locali** (su un record alla volta) e **globali**.

# Organizzazione logica

## - organizzazione non sequenziale (hash)

Questo tipo di organizzazione è un'evoluzione dell'organizzazione relativa e presenta il vantaggio del riconoscimento di un record in base alla chiave anziché alla posizione.

Vedremo più avanti ulteriori dettagli di questo tipo di organizzazione.

## Argomenti

- Organizzazione ed accesso
- Organizzazione
- Organizzazione fisica
  - Supporti ad organizzazione sequenziale
  - Supporti ad organizzazione casuale
- Organizzazione logica
  - Organizzazione sequenziale
  - Organizzazione non sequenziale

## 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