

Il file system e l'architettura delle informazioni

(Nota: una discussione sull'informazione e sul suo significato è presente sullo Sciuto Introduzione e Capitolo 3. I concetti di file e file system sono definiti sullo Sciuto Cap. 5 Pag. 212)

Di P. Foglia

1.0 le informazioni ed i supporti per le informazioni

L'ACM ha definito l'informatica come "lo studio sistematico degli algoritmi che descrivono e trasformano l'informazione: la loro teoria, analisi, progetto, efficienza, realizzazione e applicazione". Per un'altra definizione, l'informatica è la scienza della rappresentazione e dell'elaborazione dell'informazione.

In entrambi le definizioni, centrale è il concetto di informazione. L'informazione è "ciò che viene trasmesso o comunicato" nell'ambito di una utilità e funzionalità pratica immediata. Fondamentale nel concetto di informazione sono i dati: i dati, infatti, possono essere definiti come gli elementi di cui si dispone per formulare un giudizio, caratterizzare un fenomeno o pervenire alla soluzione di un problema. I dati sono perciò l'oggetto dell'informazione. Tramite i dati (e dunque l'informazione che essi contengono), è possibile ragionare e comunicare sulle cose senza averle fisicamente presenti.

Supponiamo ad esempio di dover far passare un mobile da una finestra. Un primo modo di risolvere il problema consiste nel prendere il mobile e fare tutti i possibili tentativi "a mano". In tal caso occorre disporre sia del mobile che della finestra. Un secondo modo consiste nel prendere le misure dei mobili, ossia dei dati e dunque delle informazioni, elaborare tali dati (prenderne la dimensione massima) e verificare che siano compatibili (inferiori in tale caso) con le dimensioni della finestra. Il vantaggio di tale approccio consiste nella possibilità di capire (ossia elaborare l'informazione) se il problema (far passare un mobile da una finestra) è risolvibile anche senza avere il mobile a disposizione. Inoltre è possibile elaborare l'informazione con strumenti che niente hanno a che fare con il problema in oggetto. Ad esempio, il problema di far passare un mobile da una finestra si risolve tramite operazioni di confronti, ossia operazioni matematiche, che non hanno niente in comune con il mondo dei mobili e delle finestre.

Concetto strettamente legato a quello di informazione è quello di supporto. I supporti sono gli strumenti che l'uomo ha creato per trattare (ossia per creare, conservare, modificare e comunicare) l'informazione. Tornando all'esempio del mobile e delle finestre, sono in grado di generare dei dati (le misure) ma come li trasmetto? Come li conservo? Ecco dunque l'esigenza di un supporto per la memorizzazione e la gestione dell'informazione. Ovviamente il supporto contiene informazione ma non è l'informazione. Esempi di supporti sono la carta, le fotografie, i dischi, i nastri, etc.

2.0 Il sistema Operativo, il file system, i file e le directory

Anche i calcolatori costituiscono dei supporti per memorizzare l'informazione. In particolare, la memoria di massa, ossia il disco, rappresenta un supporto per memorizzare l'informazione. Si è visto che un disco è organizzato in blocchi, settori, tracce. E dunque l'informazione nel disco, a livello fisico, viene memorizzata in tali blocchi, settori e tracce. D'altra parte, affinché il trattamento da parte dell'uomo dell'informazione (che è l'unico a comprenderne il "significato") contenuta nella memoria di massa avvenga con la stessa semplicità con cui l'uomo manipola solitamente l'informazione, ad esempio leggendo documenti, scrivendone altri, raggruppandoli in cartelle, è necessario che il "calcolatore" dia all'uomo degli strumenti che non siano semplicemente i blocchi

ed i settori. Il sistema operativo è il componente del calcolatore (inteso come insieme di hw più sw) che fornisce tale funzionalità. In particolare quel componente del sistema operativo che è **il file system**.

Quali strumenti il file system mette a disposizione dell'utente per gestire le informazioni?

Lo strumento di base è il file.

Il **file** è un contenitore logico caratterizzato da un **nome** (filename), un **contenuto** (che rappresenta l'informazione che voglio memorizzare in esso) e degli **attributi** (la dimensione, la data di creazione, la data di modifica, i diritti). Il file può essere pensato come l'equivalente di un documento scritto su carta. Il nome del documento è il nome del file, il contenuto del documento (l'insieme di parole, grafici, etc ossia di simboli contenuti nel documento) è il contenuto del file. Ovviamente non è detto che il documento, per il fatto di avere un contenuto, contenga anche informazione. Ciò è vero solo se i simboli contenuti al suo interno hanno , per noi lettori, un significato. Ciò vale tanto per il documento su carta quanto per il documento in formato elettronico.

Se il file system fosse organizzato come un unico contenitore di file, non sarebbe possibile creare due file con lo stesso nome e la gestione diventerebbe difficoltosa nel momento in cui ho molti file. Pertanto il file system introduce anche lo strumento delle directory o cataloghi o direttori o cartelle. Una **directory** è un contenitore logico che può contenere file ma anche altre cartelle. Inoltre ogni file system prevede che tutte le cartelle siano contenute in un'unica directory che prende il nome di radice o root. Tale radice è unica nel sistema operativo Unix, è una per ogni dispositivo nel sistema operativo Windows.

Come individuare ogni file all'interno del file-system?

Il nome del file di per se non basta, perché posso avere più file con lo stesso nome. Il file si identifica univocamente tramite un nome completo o percorso assoluto (**pathname assoluto**), costruito scrivendo ordinatamente i nomi dei vari nodi che si incontrano a partire dalla radice, intervallati da un carattere di separazione (\ in windows, / in Unix).

Durante le operazioni di gestione del file system, in ogni istante è definita una **directory corrente**, che rappresenta la directory in cui l'utente attuale sta lavorando. Ogni file può allora essere identificato, oltre che dal pathname assoluto, anche dal **pathname relativo** alla directory corrente. Nella specifica dei pathname relativi, la directory . rappresenta la directory corrente e la directory .. rappresenta la directory immediatamente superiore a questa.

3.0 I servizi di base del file system

I servizi di base offerti dal file system per la manipolazione dei file prevedono:

1. la visualizzazione del contenuto di un file e di una directory
2. la creazione di file e directory
3. la possibilità di navigare fra file e directory
4. la possibilità di cancellare/copiare/spostare file e directory

Tali servizi sono offerto o tramite dei comandi dati ad un interprete dei comandi (una shell) o tramite una interfaccia grafica.

Un ulteriore servizio è quello di collegamento o link o alias o shortcut. Il collegamento è un riferimento ad un altro oggetto presente nel file system (sia esso file o directory). Ciò consente di avere lo stesso file presente in più directory, ma di risparmiare in termini di occupazione del file, in

quanto non è necessario effettuare il duplicato del file. Le modifiche che faccio in un file si propagano anche al secondo (perché in realtà sono lo stesso file). L'aggiornamento delle copie multiple è invece un problema che riguarda la duplicazione dei file.

Per i sistemi operativi multiutente, inoltre, è possibile effettuare il controllo degli accessi ai file. Il file system allora offre anche comandi per:

- 1) visualizzare i permessi di accesso ai file.
- 2) modificare tali permessi.

4.0 L'organizzazione delle informazioni

Quando si dispone di grosse quantità di informazione, è indispensabile organizzare con un qualche criterio tale informazione. poichè l'informazione è qualcosa di astratto, sono i supporti che l'uomo ha inventato per crearla, conservarla, modificarla, e comunicarla che devono essere organizzati. Ciò vale anche per le informazioni contenute nella memoria di massa di un computer, e dunque gestite dal file system.

Obiettivo di una qualunque tecnica di organizzazione e gestione consiste nel facilitare le operazioni di:

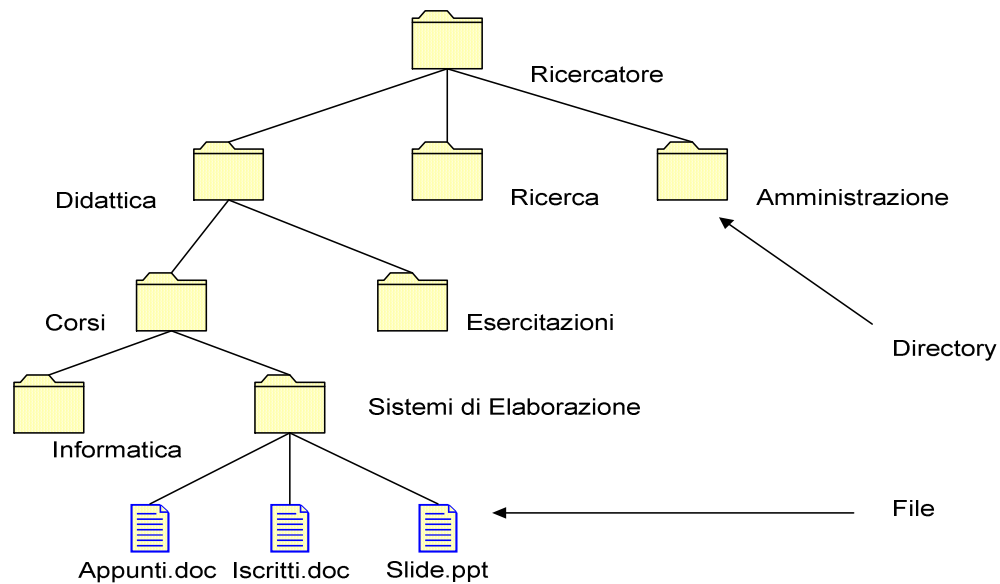
- 1) recupero dell'informazione
- 2) eliminazione dell'informazione
- 3) modifica/aggiornamento dell'informazione
- 4) copia dell'informazione

Consideriamo ad esempio ad una libreria. La libreria contiene dei libri; i libri sono uno strumento per memorizzare l'informazione. Senza una qualche organizzazione, sarebbe impossibile gestire la libreria con una qualche efficienza, e dunque l'informazione in essa contenuta. Supponiamo di disporre i libri "a caso". Sarebbe difficoltoso indicare ad un cliente con una richiesta precisa la collocazione di un libro, ma sarebbe difficoltoso anche indicare ad un cliente con una richiesta di genere ("vorrei un romanzo") come muoversi, e sarebbe difficoltoso per lui visionare tutti i romanzi presenti nella biblioteca. Sarebbe inoltre difficoltoso verificare la disponibilità di un testo, o l'insieme dell'offerta di un genere, per effettuare degli ordini periodici. Una buona organizzazione consisterebbe nel "classificare" i libri, e disporli sulla base di tale classificazione: ad esempio i romanzi, i libri di storia, di attualità, di scienza, etc. La classificazione potrebbe non necessariamente essere "ad intersezione vuota": si potrebbe avere, oltre alle varie area tematiche, anche l'area dei best-seller, che includerebbe elementi di tutte le classi. Inoltre a tale disposizione "concettuale" si potrebbe far coincidere anche un disposizione fisica: dunque organizzare lo spazio fisico della libreria in "aree": l'area di storia, l'area di geografia, l'area di scienze; in ogni area collocare degli scaffali, ossia dei contenitori di libri, ed anche per gli scaffali usare un'organizzazione: libri storici della preistoria, dell'epoca romana, del medioevo, etc. I problemi di una disposizione "a caso" verrebbero in questo modo superati.

La stessa tecnica può essere utilizzata nel caso dell'informazione memorizzata in un calcolatore, e dunque si può adottare un'organizzazione per i file contenuti nel file system. In tale caso, non si dispone di tanti strumenti di "disposizione fisica" (aree, scaffali, cesti, etc) ma solo di due strumenti: i file e le directory. Tuttavia, dando dei nomi adeguati, è possibile ottenere delle organizzazioni molto sofisticate.

Pensiamo ad esempio all'insieme dei documenti elettronici di un ricercatore. Questi riguarderanno la didattica, la ricerca ed i documenti amministrativi. I documenti della didattica riguarderanno i

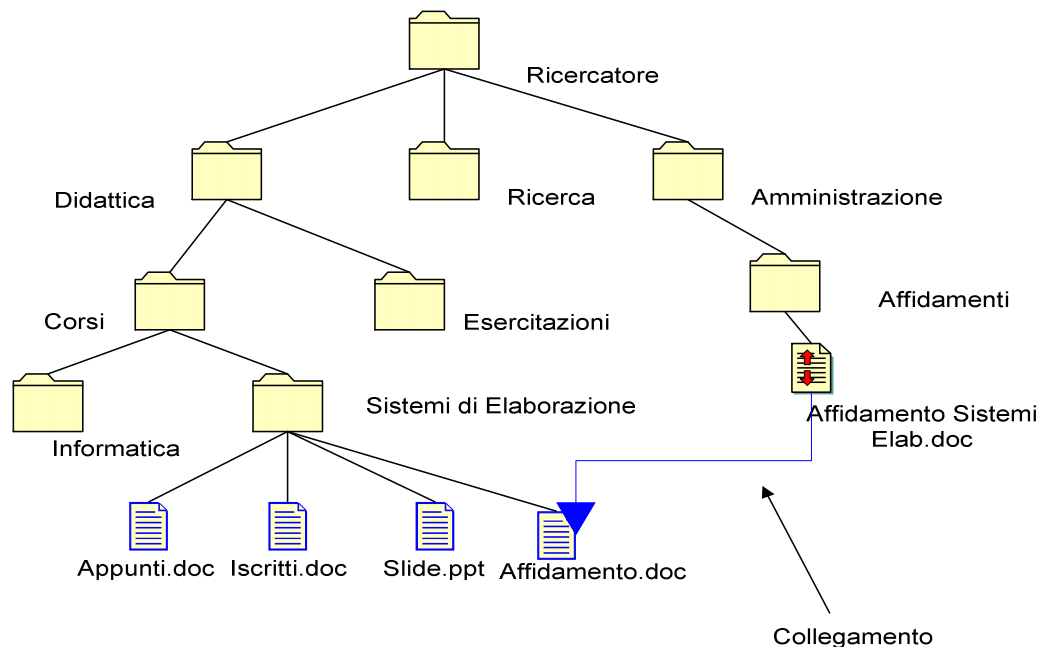
corsi e le esercitazioni ed includeranno le presentazioni, gli appunti, gli iscritti. Un possibile modo per organizzare, all'interno del file system, tale informazione, sarà il seguente:



L'utilizzo dei link.

Supponiamo adesso di avere, fra i documenti del corso, anche i contratti di affidamento dell'insegnamento. Tali documenti fanno parte delle informazioni del corso, ma sono anche documenti amministrativi. Come rappresentare tale situazione?

Le soluzioni sono due: in prima battuta si potrebbe pensare di duplicare il file "affidamento". In tal caso però, si pone il problema della gestione delle copie: se devo effettuare una modifica al documento e non ricordo di avere due copie, avrò le due copie non più consistenti. In alternativa, ricorro ad un link.



Commenti alla funzione di ricerca

I sistemi operativi forniscono solitamente anche un comando per effettuare la ricerca. Si potrebbe obiettare allora che l'organizzazione è inutile per ricercare file, dal momento che con il comando "cerca" si riesce comunque a trovare il file. Il problema di tale comando è che occorre ricordare informazioni sul file, in particolare il nome (o qualche altro dato, come l'estensione del file). Tale informazione è abbastanza "puntuale", ossia precisa.

Con l'organizzazione proposta nell'esempio precedente, dovendo individuare un documento di ricerca, ossia avendo informazioni solo legate allo scopo del file (e non puntuali) riesco ad adottare una strategia per la ricerca del documento (esploro il contenuto di tutte le cartelle sotto la directory "ricerca").

Esercitazione:

- 1) Organizzare il file system in modo che posso consentire la gestione dei documenti della facoltà di ingegneria. La Facoltà di Ingegneria è organizzata in Dipartimenti; in particolare sono presenti i dipartimenti di Ingegneria dell'Informazione, di Ingegneria Elettrica ed Ingegneria Meccanica. Ad ogni dipartimento afferisce una struttura di ricerca, un corso di laurea ed un corso di laurea specialistica. I corsi di laurea (entrambi) riguardano l'attività didattica di un Dipartimento. Inoltre esiste la Segreteria Didattica della Presidenza che, essendo la coordinatrice di tutta l'attività didattica della facoltà, deve disporre di tutti i documenti di tutte le strutture didattiche presenti in facoltà.
- 2) Inserire in tale struttura il documento word "iscritti al corso di laurea in ingegneria gestionale", il rapporto di ricerca "progetto per un calcolatore SPP" che è un documento .txt, la relazione del Preside sulla didattica in formato word.
- 3) Essendo il Preside, verificare "a mano" che tutti i dipartimenti abbiano preparato la lista degli iscritti ai corsi di laurea triennale.
- 4) Andare sul Desktop. Creare il documento "iscritti al corso di laurea specialistica in Ingegneria Informatica". Inserire il documento nella directory di pertinenza utilizzando dei comandi di shell ed i pathname assoluti. Creare il documento "iscritti al corso di laurea specialistica in Ingegneria Meccanica". Inserire il documento nella directory di pertinenza utilizzando dei comandi di shell ed i pathname relativi.
- 5) Creare uno script che consenta di "creare" un nuovo dipartimento. Come è possibile renderlo parametrico?