

Fondamenti di informatica

A cura di Pierpaolo N.

In questa guida vengono spiegate le basi dell'informatica partendo dalla macchina di Turing fino ad arrivare alla tecnologia dei giorni nostri. Venne realizzata durante il secondo anno accademico della facoltà di Ing. Elettrica presso il Politecnico di Bari.

Nb: la suddivisione in paragrafi e capitoli, è stata utilizzata dall'autore come riferimento al testo universitario; pertanto tali suddivisioni non fanno riferimento alla struttura data alla guida.

Nella speranza di averVi fatta cosa gradita, Vi ringrazio della preferenza.

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presepando> nella home page della sezione dedicata a Melfi.

NOZIONI INTRODUTTIVE

CENNI STORICI

I sistemi di numerazione nacquero molto presto nella storia dell'umanità, ve ne sono tracce di diversi tipi: gli Egizi dal 3000 a.C. utilizzavano una numerazione basata sui simboli 1, 10, 100, 1000 ..., i Maya utilizzavano sistemi di numerazione in base 18 e 20, i Babilonesi uno in base 10 (per gli scambi commerciali) e uno in base 60 (per i problemi scientifici); altre tracce di sistemi in base 3,4,5,6,8 e 20 sono stati trovati presso gli Indiani dell'America del Nord.

Oggi se pur la numerazione in base 10 ha prevalsa, per la sua praticità sulle altre, si usano ancora sistemi di numerazione binaria (2), ottale (8) ed esadecimale (16) perché presentano, per i calcolatori elettronici, caratteristiche più vantaggiose.

Il sistema di numerazione attuale è una evoluzione del sistema di numerazione utilizzato in India nel III secolo a.C.; tale sistema, infatti, non comprendeva l'uso del simbolo zero che, con la sua introduzione, ha reso possibile la realizzazione di un vero sistema posizionale dove il valore della cifra dipende dalla posizione.

Nell' VIII secolo d.C. il matematico arabo Muhammed ibn-Musà, nato a Khavarezem e noto col nome al-Khuwarismi, scrisse dei trattati, poi tradotti in latino, in cui illustrò i procedimenti dell'aritmetica decimale; mentre nel XVI secolo d.C. venne introdotto l'uso della virgola, o punto decimale, che permise di estendere il sistema decimale oltre il campo dei numeri interi.

Il sistema di numerazione binario fece la sua prima apparizione in Europa verso il 1600 su alcuni manoscritti di T. Harriot; le regole di tale numerazione furono poi esposte da G.W. Leibniz all'Accademie des Sciences nel 1703.

Questo sistema si diffuse rapidamente a causa dell'aumento di interesse per la costruzioni di apparecchiature capaci di svolgere automaticamente delle operazioni aritmetiche. I simboli usati, detti bit (**BI**nary **digi**T), sono 0 e 1.

Un grande vantaggio di tale sistema è costituito dal piccolo numero di simboli grafici fondamentali, mentre un grande svantaggio è costituito dal grande numero di cifre binarie necessarie per rappresentare un numero, rispetto a quelli necessari negli altri sistemi di numerazione ($2^n - 1$: numero massimo realizzabile con n bit, $2^{n-1} - 1$: numero minimo realizzabile con n bit).

A causa dello svantaggio del sistema binario ed alla scarsa intelligibilità per l'uomo, si adottano dei sistemi di numerazione capaci di ridurre questi inconvenienti ma di facile conversione da e verso il sistema binario; essi sono il sistema di numerazione ottale ed il sistema di numerazione esadecimale. Il sistema ottale è caratterizzato dall'estrema facilità di conversione da e verso il sistema di numerazione binario.

Nella numerazione esadecimale la conversione è simile a quella fatta per il sistema ottale, in questo caso i gruppi devono essere formati da quattro cifre partendo sempre dal punto esadecimale 0, i sua mancanza, dalla cifra di destra.

NOZIONI DI ALGEBRA BOOLEANA.

Il funzionamento di un calcolatore elettronico si basa sull'impiego di elementi aventi due stati di funzionamento, sia per l'economicità di tali circuiti sia per la possibilità di esprimere u qualsiasi dato mediante una sequenza di 0 e 1.

Le nozioni fondamentali dell'algebra che permettono di trattare adeguatamente questi simboli binari, prendono il nome di algebra booleana in onore al matematico George Boole (1815-1864) che per primo, nel 1854, ne pubblicò i concetti generali.

POSTULATI PAG 29 DEL LIBRO

TEOREMI FONDAMENTALI PAG 33 DEL LIBRO

Nei calcolatori elettronici le operazioni booleane sono realizzati da circuiti detti blocchi logici. Essi si compongono di diversi componenti come: diodi, resistenze, transistor...; per realizzare una qualsiasi funzione booleana si fa uso di reti logiche realizzate tramite uno opportuno collegamento dei blocchi logici elementari che la costituiscono.

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presependo> nella home page della sezione dedicata a Melfi.

Le operazioni booleane di somma e prodotto, sono svolte rispettivamente dalla porta logica OR (detta anche blocco logico elementare) a due ingressi ed una sola uscita, e dalla porta logica AND. La realizzazione di un complemento si ottiene con l'uso del blocco logico elementare NOT.

A seconda della tecnologia adoperata, può essere utilizzato un blocco logico NAND per sostituire una AND ed una NOT.

Per la realizzazione di circuiti logici economici e per la loro minimizzazione, si tende a semplificare la funzione booleana che si vuole ottenere tramite le mappe di Karnaugh e i teoremi di De Morgan.

Esistono due diversi tipi di reti logiche, essi sono:

- Reti logiche combinatorie;
- Reti logiche sequenziali.

Nelle reti logiche combinatorie i valori delle uscite sono determinate univocamente dalla configurazione degli stati degli ingressi; il loro comportamento non dipende quindi né dal tempo né da quanto è successo prima, ovvero non dipendono dalla storia della rete.

Nelle reti sequenziali i valori delle uscite dipendono sia dagli stati degli ingressi sia dalla storia della rete; per tale motivo si rende necessaria la presenza di elementi di memoria capaci di conservare il valore assunto dalla variabile booleana: l'elemento più noto è il flip-flop (detto circuito distabile).

CODIFICAZIONE DEI DATI.

La codificazione dei dati è un processo che permette di stabilire una corrispondenza biunivoca tra due diverse simbologie di simboli. Condizione necessaria e sufficiente affinché si parli di codifica:

1. ogni simbolo deve essere codificato in numero distribuito;
2. ogni codice deve avere una sezione iniziale diversa.

Nel codificare dei dati con l'alfabeto binario (0,1), si possono incontrare delle difficoltà in lettura e scrittura; per evitare ciò si utilizzano altri tipi di codici. Uno dei più diffusi sistemi di codificazione numerica è quello detto BCD (**B**inary **C**oded **D**ecimal) in cui ogni cifra decimale, del numero da codificare, è rappresentata tramite 4 bit che esprimono il valore binario; es. codifica di 394: 0011 (3) 1001 (9) 0100 (4).

Questo sistema, detto pesato perché il valore di ogni cifra decimale viene ottenuto mediante somma pesata delle 4 cifre binarie che la rappresentano, presenta delle complicazioni e difficoltà nell'eseguire addizioni con somme maggiori di 9 e nella complementazione.

Questi due ultimi problemi vengono superati con il **codice di accesso 3** che, basandosi sul codice BCD, aggiunge un 3 binario alla rappresentazione ciascuna cifra decimale.

Nella codificazione di caratteri alfa numerici si distingue un alfabeto interno generalmente composto da 64 simboli:

- 26 lettere maiuscole dell'alfabeto;
- 10 cifre decimali;
- 28 caratteri vari.

Per la loro codificazione sono stati inventati codici a 6 bit (64 simboli) usati per sostituire quelli a 4 bit (16 simboli).

Allo scopo di permettere una estensione all'insieme di caratteri sono nati codici a 7 bit detti ISO (**I**nternational **S**tandard **O**rganization) in grado di trattare 128 caratteri, e codici a 8 bit.

Questi ultimi si dividono in:

- codice EBCDIC (**E**xtended **B**inary **C**oded **D**ecimal **I**nterchange **C**ode) utilizzato nei calcolatori IBM;
- codice ASCII (**A**merican **S**tandard **C**ode for **I**nformation **I**nterchange).

In questi codici il gruppo di 8 bit è detto byte.

Nella codificazione dei numeri segnati, dove ad ogni cifra si associa un numero fisso di bit, al segno + o - si attribuisce un bit posizionato alla sinistra della cifra più significativa. Nel caso in cui le operazioni aritmetiche dovessero generare un riporto che superi il numero di bit prefissato, si ha un traboccamento (overflow) e il risultato viene troncato nella sua cifra più significativa.

La rappresentazione in virgola mobile è usata nei problemi tecnici o scientifici; essa può generare un traboccamento quando il valore dell'esponente, di un numero, è maggiore di quello prefissato (overflow) o quando è più piccolo di quello prefissato (underflow).

ALGORITMI.

GENERALITÀ SUGLI ALGORITMI.

Il termine algoritmo scambiato per una stropicatura di logaritmo, deriva dal paese natale, al-Khuwarismi, del matematico arabo Muhammed ibn-Musà e ad esso si associava il significato di procedimento di calcolo. Oggi si è avuto un vero e proprio sviluppo della teoria degli algoritmi per cui al termine algoritmo viene associata la seguente definizione:

l'algoritmo è un insieme di istruzioni che definiscono una sequenza di operazioni atta a risolvere una classe di problemi.

Alla definizione data si associano tre importanti caratteristiche di algoritmo, esse sono:

- carattere finito: il numero di istruzioni di cui si compone un algoritmo deve essere finito;
- carattere deterministico: l'algoritmo non deve essere ambiguo ed eseguito da persone o calcolatori differenti deve dare sempre lo stesso risultato;
- carattere di realizzazione pratica: tutte le operazioni di cui si compone un algoritmo devono essere effettivamente eseguibili.

L'esigenza di avere una descrizione chiara e dettagliata delle sequenze operative di un algoritmo, ha portato alla realizzazione di una rappresentazione grafica conosciuta come: diagramma logico, diagramma a blocchi, flow-chart.

Esso è composto da vari blocchi, collegati da frecce, aventi varie forme a seconda delle operazioni da eseguire.

La realizzazione di un programma avviene, generalmente, tramite la programmazione strutturata; essa è caratterizzata dalle seguenti regole:

- costruzione di algoritmi dal generale al particolare per raffinamenti successivi (top-down);
- uso disciplinato delle strutture di controllo.

Col procedimento top-down vengono descritti in modo preciso e formale i diversi livelli dell'algoritmo sino ad ottenere una descrizione dettagliata pronta ad essere tradotta in un linguaggio specifico. Se questo procedimento è ben fatto si riducono i problemi fra le varie parti del programma e si agevola la manutenzione.

Col termine strutture di controllo si intende:

- diramazione semplice (then);
- diramazione multipla;
- iterazione (do);
- iterazione a condizione finale (repeat);
- iterazione a contatore;
- struttura loop-exit.

MACCHINA DI TURING.

La macchina di Turing (matematico inglese) si compone di:

- una unità di memoria esterna;
- una unità di memoria interna;
- una unità di controllo;
- una unità logica.

L'unità di memoria esterna consiste in un nastro illimitato e suddiviso in sezioni, dette celle, nelle quali è contenuto l'alfabeto esterno. L'unità di memoria interna ha il compito di immagazzinare i simboli dell'alfabeto interno.

L'unità di controllo ha il compito di analizzare le celle e di gestire l'avanzamento del nastro; ad essa si collega una memoria di controllo che ha il compito di memorizzare uno dei tre simboli:

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presependo> nella home page della sezione dedicata a Melfi.

S,D,F per “ricordare” gli ordini dati dall’unità di controllo dove i simboli hanno il seguente significato:

- S: esaminare cella immediatamente a sinistra;
- D: esaminare cella immediatamente a destra;
- F: esaminare la stessa cella.

L’unità logica, avente due ingressi e tre uscite, analizza i dati di ingresso, uno dell’alfabeto esterno (cella esaminata) e uno di quello interno, e fornisce in uscita una terna avente tre simboli, essi sono:

- Un simbolo dell’alfabeto esterno;
- Un simbolo dell’alfabeto interno;
- Uno dei tre simboli di movimento (S,D,F).

L’unità logica può quindi essere considerata come una funzione rappresentabile tramite una tabella detta matrice funzionale.

La macchina di Turing per il suo funzionamento necessita di una particolare condizione iniziale, ovvero:

- In un numero finito di celle devono essere contenuti i simboli dell’alfabeto esterno diversi dal simbolo vuoto;
- L’unità di controllo deve esaminare una determinata cella;
- La memoria interna deve contenere un simbolo di stato diverso dal quello di arresto.

Quando sussiste questa condizione la macchina funziona nel seguente modo:

- l’unità di controllo preleva il simbolo s , contenuto nella cella esaminata e lo invia all’unità logica;
- la memoria interna invia all’unità logica il simbolo di stato q_1 ;
- l’unità logica, in funzione della coppia di simboli in ingresso, fornisce la terna di simboli s_j , q_j , m_j , definiti dalla matrice funzionale;
- il simbolo s , viene registrato nella cella esaminata in sostituzione di quello precedentemente contenuto;
- il simbolo di stato q_j viene registrato nella memoria interna al posto del simbolo precedentemente contenuto;
- il simbolo di movimento m_j viene registrato nella memoria di controllo e l’unità di controllo passa ad esaminare la cella immediatamente a sinistra, immediatamente a destra, o esamina nuovamente la stessa cella, a seconda del valore di m_j

Per ipotesi fondamentale della teoria degli algoritmi si intende che un qualsiasi algoritmo può essere realizzato tramite la macchina di Turing.

STRUTTURA DI UN CALCOLATORE ELETTRONICO

ANALISI DEL PROCESSO DI ELABORAZIONE AUTOMATICA

Il processo di elaborazione automatico è caratterizzato da alcune funzioni elementari:

- Funzione di introduzione dati: vengono introdotti nel calcolatore, tramite apparecchiature di input, dei dati da analizzare;
- Funzione di estrazione dati: i dati elaborati vengono estratti dall’elaboratore tramite delle apparecchiature di output;
- Funzione di memorizzazione dati: vengono memorizzati, su apposite memorie, i dati iniziali, l’elaborazione eseguita e i dati finali.
- Funzione di calcolo: nel caso in cui l’algoritmo preveda lo svolgimento di operazioni logiche o aritmetiche, esse saranno svolte dall’unità logica e aritmetica;
- Funzione di governo: viene svolta dall’unità di governo che gestisce l’elaborazione dati e tutti quei componenti che svolgono l’operazione stessa.

Generalmente l’insieme costituito dall’unità di governo, dalle memorie e dall’unità aritmetica e logica è detto unità centrale (Central Processing Unit).

Le apparecchiature di introduzione ed estrazione dati sono molto diversificate tra loro, le più diffuse sono:

- Scheda perforata;

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all’indirizzo: <http://utenti.quipo.it/presepano> nella home page della sezione dedicata a Melfi.

- Nastro perforato;
- Lettura ottica;
- Lettura magnetica;
- Stampanti.

La scheda perforata, se pur soggetta ad un rapido declino, è stata per lungo tempo il principale strumento di input per i calcolatori elettronici

Questo supporto è di cartoncino avente dimensioni standard, un buon isolamento elettrico ed una discreta resistenza meccanica. Essa viene divisa in 80 colonne, numerate sequenzialmente da 1 a 80, e 12 righe. Le righe sono soggette ad una numerazione particolare: le prime due, non numerate, sono dette *righe sopraelevate* e, convenzionalmente, si indicano con i numeri 11 e 12, mentre le restanti righe, dette *normali*, vengono numerate da 0 a 9.

I dati vengono registrati su queste schede tramite appositi dispositivi detti perforatori di schede, e i fori (rettangolari) sono realizzati tramite punzoni. I dispositivi atti alla lettura sono detti lettori di schede e si dividono in :

- seriali;
- paralleli.

I primi sono più economici, e sono caratterizzati da una lentezza maggiore rispetto ai secondi; la lettura dei dati avviene colonna per colonna con l'invio di 12 segnali per volta. I lettori di schede a funzionamento parallelo sono più costosi dei primi ma caratterizzati da una maggiore velocità di lettura; la lettura dei dati avviene con l'analisi di tutte le 80 colonne e l'invio dei dati è contemporaneo.

Il sistema di lettura può essere di tipo elettromeccanico o fotoelettrico.

Entrambe i sistemi basano il loro principio di funzionamento sull'invio di tensioni al calcolatore; nel primo sistema ciò avviene tramite rullo (sottotensione) e spazzolini, mentre nel secondo sistema avviene per mezzo di una fonte luminosa e delle cellule fotosensibili.

La codifica dei segnali avviene mediante delle tabelle che costituiscono il codice di perforazione (il più noto è il codice di HOLLERITH), esse stabiliscono una corrispondenza tra le perforazioni delle schede e i simboli dell'alfabeto.

Un altro supporto tradizionale è costituito dal nastro perforato. Esso è una striscia continua, detta anche banda, di carta opaca avvolta su bobine e di lunghezza variabile. Si suddivide in righe perpendicolari all'asse sul quale si realizzano le perforazioni; i fori più piccoli servono per il trascinarsi ed il posizionamento esatto del nastro.

Esistono diversi tipi di nastro perforato, questa distinzione è dovuta a:

- forma dei fori: tondi o rettangolari;
- codice di perforazione utilizzato;
- numero massimo di perforazioni su una stessa riga.

L'ultima caratteristica determina il numero massimo di caratteri rappresentabili per riga. In genere i nastri vanno da 5 tracce a 8 tracce, ovvero da $2^5=32$ a $2^8=256$ caratteri diversi.

Nel caso di nastro perforato a 8 tracce, una traccia è utilizzata a scopo di controllo; tale traccia, detta di *parità* o di *disparità*, ha il compito di rendere sempre pari, ovvero sempre dispari, il numero delle perforazioni relative ad ogni carattere. Questo metodo consente una più facile rilevazione di errori, una più facile correzione (effettuata sulla traccia di controllo); l'unico inconveniente è che i caratteri rappresentabili scendono da 2^8 a 2^7 , ovvero da 256 a 128.

I dispositivi di lettura e scrittura sono simili a quelli utilizzati per le schede perforate; per il nastro perforato è poco usato il metodo di lettura elettromeccanico perché è causa di usura del nastro stesso. È da notare che sia la lettura che la perforazione del nastro può avvenire solo in modo seriale!!!

La lettura magnetica è principalmente utilizzata nel campo bancario per il trattamento automatico di assegni e cambiali. Questa tipologia di lettura, detta MICR (**M**agnetic **I**nk **C**haracter **R**ecognition), si basa sull'esplorazione della superficie dei supporti e sulla rilevazione della presenza, o meno, di caratteri convenzionali.

I caratteri sono impressi su un supporto con dell'inchiostro contenente materiale elettromagnetico. Prima della lettura dei caratteri, si procede ad una magnetizzazione, tramite

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presepano> nella home page della sezione dedicata a Melfi.

magnete permanente, dei caratteri stessi in modo tale da far magnetizzare l'ossido di ferro contenuto nell'inchiostro. Dopo tale operazione l'ossido di ferro genera un campo magnetico, legato alla forma del carattere, in grado di essere riconosciuto dalla testina di lettura.

I più diffusi tipi di caratteri magnetici sono: E13B, utilizzato in America, e il CMC7, utilizzato in Europa, detto anche carattere magnetico codificato (Caractère Magnétique codé 7).

Il sistema di lettura ottica, detto anche OCR (**O**ptical **C**haracter **R**ecognition) interpreta i simboli grafici manoscritti, dattiloscritti, stampati aventi forma tradizionale.

Con la lettura ottica, l'obiettivo che si vuole raggiungere è quello di utilizzare il documento originale come supporto di input per il calcolatore. Le varie tecniche per effettuare tale lettura prevedono, come elementi comuni, dispositivi fotosensibili in grado di rilevare dei segnali ottici e convertirli in impulsi elettrici. I metodi di rilevamento più diffusi sono:

- rilevamento a scansione;
- rilevamento a lettura analogica;
- rilevamento a matrice bidimensionale;
- rilevamento morfotopologico.

Nei primi due casi, in generale, il rilevamento si realizza tramite un raggio luminoso che analizza la zona contenente i caratteri. Il raggio di luce, proiettato sulla faccia del documento, diffonde più o meno luce, in base alla quantità di inchiostro, la quale è convertita, tramite fotomoltiplicatore, nel corrispondente segnale elettrico.

Negli ultimi due casi, generalizzando, si ha un confronto tra i caratteri rilevati e i caratteri campioni presenti nella macchina.

L'inconveniente del sistema ottico è dovuto alla sensibilità dei dispositivi utilizzati ed agli eventuali segni estranei al testo che, per quanto riguarda la lettura magnetica, non influenzano il riconoscimento poiché, non essendo di materiale ferromagnetico, ne altera né genera un campo magnetico.

Un altro metodo di input, in grande diffusione, è il codice a barre. Esso si compone di una successione di barre scure, divise da spazi bianchi, aventi diversi spessori e rappresentano, mediante codice binario, cifre decimali e, in qualche raro caso, le lettere.

Il codice a barre più diffuso è l'EAN (**E**uropean **A**rticle **N**umbering) ed è generalmente usato per marcare i prodotti alimentari. Esso deriva dal codice americano UPC (**U**niversal **P**roduct **C**ode). Con i progressivi assestamenti, a cui il codice a barre va incontro, si è avuto il codice EAN13 capace di rappresentare 12 cifre decimali, e una non codificata, con le quali vengono codificate la nazione di origine del prodotto, l'azienda produttrice ed il prodotto stesso.

La marcatura si compone di 95 elementi base, essi sono:

- 3 elementi base fissi (101): primo END MARK;
- 42 elementi base interpretati come 6 cifre decimali (7 elementi base per ogni cifra);
- 5 elementi base fissi (01010): CENTRE MARK;
- 42 elementi base interpretati come 6 cifre decimali (7 elementi base per ogni cifra) di cui l'ultima funge da cifra di controllo calcolata a partire dalle altre;
- 3 elementi fissi base (101): secondo END MARK.

Gli END MARK ed il CENTRE MARK sono stampati con una lunghezza maggiore degli altri elementi del codice a barre.

Un altro dispositivo di input è costituita dalla tastiera che prende il nome di apparecchiatura terminale. Essa permette l'inserimento dei dati tramite una pressione su appositi tasti alfanumerici; in genere ad essa si associa un terminale video ed una stampante.

I video permettono di vedere i dati introdotti dall'utente, a seconda di ciò che il video è capace di riprodurre si parla di:

- video alfanumerico: riproduzione di numeri ;
- video grafico: riproduzione di immagini, grafici ...
- video a cristalli liquidi.

Le figure rappresentate sullo schermo, per quanto riguarda i primi due video citati, sono ottenute dall'eccitazione dei fosfori, sostanze fotosensibili poste sulla faccia esterna dell'estremità del tubo a

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presependo> nella home page della sezione dedicata a Melfi.

raggi catodici (TRC), da arte di pennelli di elettroni opportunamente deflessi tramite campi magnetici.

Nei video a cristalli liquidi, detti LCD (**L**iquid **C**ristal **D**isplay), l'immagine è ottenuta dalla polarizzazione del cristallo liquido, contenuto nell'intercapedine delle due lastre, dovuta al campo elettrico generato da elettrodi trasparenti situati sulle lastre; le varie parti dello schermo si troveranno a tensioni diverse a seconda dell'immagine generata.

Le stampanti sono dispositivi di output che hanno il compito di stampare i risultati delle elaborazioni in modo tale che siano direttamente utilizzabili per l'uomo, esse si dividono per prestazioni e tecnica di stampa; con la prima classificazione si ha:

- stampanti seriali: stampa di un solo carattere per volta;
- stampanti parallele: stampa contemporanea dei caratteri di una riga;
- stampanti a pagina: stampa di una intera pagina;

con la seconda classificazione si ha:

- stampanti a impatto: stampa con urto meccanico su nastro inchiostro;
- stampanti non a impatto: stampa con tecniche fisiche, o chimiche, prive di urti meccanici.

Dall'unione di queste due categorie nascono le seguenti tipologie di stampanti:

- seriale a impatto;
- parallele a impatto;
- stampante non a impatto con stampa termica;
- stampante non a impatto con stampa per erosione;
- stampante non a impatto con stampa a getto di inchiostro;
- stampante non a impatto con stampa xerografica.

Nella stampante seriale ad impatto, la stampa può avvenire in due modi differenti; un primo modo è quello della stampa per punti, con il quale l'immagine è generata dall'impatto degli aghi su di un nastro inchiostro. L'altro modo è quello della stampa ad immagine intera, col quale la stampa è generata dall'impatto delle testine, su cui sono conati i caratteri, su di un nastro inchiostro; le testine possono avere varie forme, le più diffuse sono: a sfera, a cilindro e a margherita.

Anche la stampante parallela a impatto presenta due tipologie differenti di stampa, in ogni caso, però, esse sono accomunate dal fatto che stampano quasi contemporaneamente i caratteri della stessa riga.

Un primo metodo è quello della stampa a rullo, il carattere da stampare è coniato sulla generatrice di un rullo d'acciaio per tutta la sua lunghezza e la stampa è dovuta alla pressione dei martelletti sul carattere da stampare. Un secondo metodo è quello della stampa a catena (o a nastro), il carattere è coniato sopra delle placchette collegate orizzontalmente; la stampa avviene con la pressione del martelletto sul carattere da stampare.

Sia il rullo che la catena girano a velocità costanti e il comando dato ai martelletti è in sincronia con essi.

La stampante termica realizza la stampa su carta sensibile al calore per mezzo di una testina costituita da più resistenze puntiformi. Nelle stampanti a erosione, la stampa è realizzata su fogli scuri ricoperti di alluminio che, a causa di un scarica elettrica, evapora per realizzare il carattere voluto. Queste due tipologie di stampanti non hanno avuto un grande sviluppo a causa delle modeste prestazioni e degli elevati costi dei fogli per la stampa.

Al contrario la stampante a gettoni inchiostro è il dispositivo di stampa più diffuso. La stampa avviene per mezzo di un ugello che emette, ad intervalli regolari, macchie di inchiostro caricate elettrostaticamente che, deflesse da un campo magnetico, raggiungono la carta formando l'immagine.

Un altro dispositivo di stampa in netta diffusione è la stampante xerografica. Per realizzare la stampa viene prima creata una immagine elettronica che, proiettata tramite un fascio laser su un cilindro ricoperto da materiale fotoconduttore caricato elettrostaticamente, si tramuta in una immagine elettrostatica a causa della neutralizzazione selettiva del cilindro per mezzo di un laser. A tal punto l'immagine, tramite uno strato di inchiostro (toner) caricato elettrostaticamente, viene effettivamente creata grazie all'inchiostro che aderisce alle parti non neutralizzate del cilindro.

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presependo> nella home page della sezione dedicata a Melfi.

MEMORIZZAZIONE DEI DATI: GENERALITÀ 3.3

Col termine memoria si intende un dispositivo capace di immagazzinare dei dati. Tra le caratteristiche fondamentali delle memorie vi sono: la capacità massima e la velocità di accesso.

La capacità è espressa in multipli del byte, come Kbyte e Mbyte, ove i simboli K e M hanno i rispettivi valori: 1024 e 1048576.

La velocità di accesso indica, invece, l'intervallo di tempo che intercorre dal momento in cui viene richiesto un dato immagazzinato e il momento in cui è disponibile. Si distinguono quindi tre categorie di memorie:

- memorie ad accesso uniforme;
- memorie ad accesso diretto;
- memorie ad accesso sequenziale.

Le prime, che costituiscono la memoria principale di un calcolatore, sono così dette perché il tempo di accesso è uguale per tutti i dati in essa immagazzinati. Le seconde, utilizzate come memorie ausiliarie, hanno un tempo di memoria variabile caratterizzato dalla posizione degli elementi, di rotazione, atti all'accesso dei dati. Le terze sono caratterizzate da supporti che vengono letti sequenzialmente; non si può quindi parlare di velocità di accesso.

FUNZIONI DI CALCOLO E DI GOVERNO 3.4

Le istruzioni sono una successione di ordini atti a descrivere un algoritmo contenuto nella memoria principale di un calcolatore. L'insieme di queste istruzioni, dette istruzioni macchina, costituiscono un programma che, a sua volta, prende il nome di linguaggio base della macchina o, più in generale, di linguaggio macchina.

Le istruzioni comuni a più calcolatori sono:

- istruzioni aritmetiche e di confronto;
- istruzioni di input-output;
- istruzioni di controllo (salto).

Le prime istruzioni permettono di verificare, avendo due dati detti A e B, quale delle seguenti relazioni è verificata: $A < B$, $A = B$, $A > B$.

Le seconde consentono il trasferimento di dati tra le unità di input, o memoria di massa, e la memoria principale e da essa ad una unità di output o a una memoria di massa.

Le terze sono comandi di salto che permettono di alterare l'ordine di esecuzione delle istruzioni che compongono il programma.

L'unità di governo ha, in generale, il compito di eseguire, per ogni istruzione in linguaggio macchina, delle sequenze di comandi elementari atti a realizzare le operazioni richieste che avvengono, per passi o per cicli macchina, sotto il controllo di un orologio interno al sistema.

Il funzionamento di una unità di governo può essere schematizzato nel seguente modo:

- prelievo, dalla memoria principale, dell'istruzione da eseguire;
- deposito, nel Registro di Istruzione Corrente, dell'istruzione da eseguire;
- decodifica dell'istruzione;
- individualizzazione, tramite codice operativo, dell'indirizzo della microistruzione da eseguire.

Quando l'orologio interno abilita l'uscita si ha l'accesso alla ROM che, a sua volta, fornisce i comandi relativi alla microstruttura e l'indirizzo della microstruttura successiva. Tali istruzioni saranno eseguite sino alla fine del programma.

L'unità logica si compone, principalmente, di circuiti come:

- rete logica di calcolo;
- uno o più registri.

La rete logica di calcolo è costituita, a sua volta, da un insieme di circuiti collegati in modo tali da poter eseguire operazioni logiche, aritmetiche e di confronto. I registri, detti anche accumulatori, sono utilizzati per contenere dei dati durante l'esecuzione di una istruzione.

CENNI SULLE ARCHITETTURE 3.5

ARCHITETTURA DI SISTEMI DI ELABORAZIONE.

L'architettura dei sistemi di elaborazione è un argomento assai vasto e diversificato. In generale è caratterizzato dal tipo di calcolatore e dalle prestazioni che permettono di ottenere.

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presependo> nella home page della sezione dedicata a Melfi.

Nei medi e grandi calcolatori si distinguono le seguenti architetture:

- a unità di canale;
- MISD;
- SIMD;
- MIMD;

mentre per i piccoli calcolatori sono diffuse le architetture a bus basate sull'impiego di microprocessori.

Le unità di canale, oltre a realizzare un collegamento fisico con l'unità centrale e le periferiche, è dotato di capacità di elaborazione autonoma; si distinguono diversi tipi di unità di canale, essi sono:

- canali multiplexer;
- canali selector;
- block multiplexer.

I primi sono usati per il collegamento di più periferiche lente e hanno il pregio di gestire contemporaneamente il trasferimento dei dati tra l'unità centrale e le periferiche (pochi caratteri per volta con frequenza bassa).

I secondi sono usati per il collegamento di più periferiche veloci per volta gestendo una periferica alla volta; mentre gli ultimi sono una via di mezzo tra i due tipi prima descritti.

I sistemi MISD, SIMD, MIMD, fanno parte di architetture dove l'esecuzione delle istruzioni non è strettamente sequenziale. Nei sistemi MISD (**M**ultiple **I**nstruction **S**tream, **S**ingle **D**ata Stream) si adotta una architettura detta **pipeline** in cui si hanno più unità aritmetiche e logiche, poste sequenzialmente, dotate di compiti specifici.

Nei sistemi SIMD (**S**ingle **I**nstruction Stream, **M**ultiple **D**ata Stream) l'architettura usata è detta **array processor**, le unità aritmetiche e logiche, in numero rilevante, svolgono le stesse operazioni ma con dati diversi.

Nei sistemi MIMD (**M**ultiple **I**nstruction Stream, **M**ultiple **D**ata Stream) si hanno un insieme di processi, detti multi processor, che operano in modo asincrono.

Architettura della rete di calcolatori

Una rete di calcolatori si realizza collegando più calcolatori, ognuno avente le proprie periferiche, posti in luoghi diversi. Il collegamento tra i vari calcolatori, detti nodi, avviene tramite linee di comunicazione private o pubbliche. Per la loro estensione si distinguono diversi tipi di reti:

- reti locali: reti private di piccole dimensioni come uffici, dette **LAN**;
- reti metropolitane: reti urbane;
- reti geografiche: reti che si estendono su interi continenti, o nazioni, dette **WAN**.

Le reti locali, che operano in modo digitale, si differenziano dalle altre per la loro topologia detta regolare; queste topologie sono dette: a stella, ad anello, a bus. Una tipologia di rete locale particolare è la **ETHERNET** che si differenzia per il tipo di collegamento e per il protocollo di rete; meccanismo che regola lo scambio di informazioni sulla rete (nella **ETHERNET** è il **CSMA/CO**) su cui si basa l'architettura della rete stessa.

INTERNET, il cui primo nome fu **ARPANET**, nacque nel 1968 per collegare 4 università americane e, successivamente, si scisse in due parti dando così origine ad una nuova rete, detta **MILNET**, di uso strettamente militare. Con l'ampliarsi delle due reti si sviluppò l'attuale rete informatica quale è **INTERNET** che, oggi, è gestita da una sola società. I protocolli di internet più diffusi sono: **http** (per pagine virtuali), **ftp** (per immagini), **www** (per pagine web).

STRUTTURE ASTRATTE DEI DATI.

PREMESSA.

Tutti i dati, presenti in un calcolatore, sono dotati di una struttura logica che viene detta struttura astratta di dati. Con questo termine si indica un insieme finito di elementi e un insieme di leggi che definiscono le relazioni tra gli elementi.

La nomenclatura in uso è molto discordante, infatti si può presentare il caso in cui una stessa struttura è indicata con più termini o che un termine venga usato per più strutture; i termini più utilizzati sono i seguenti:

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presependo> nella home page della sezione dedicata a Melfi.

- lista;
- coda, pila, doppia coda;
- matrice;
- tavola;
- grafo;
- albero.

Con il termine lista si indica un insieme finito di dati in cui esiste una legge, di ordinamento, che stabilisce una corrispondenza biunivoca tra i suoi elementi e l'insieme dei primi n numeri naturali. In questo tipo di struttura si può facilmente stabilire quale elemento sia primo o ultimo, e quale tra due elementi, precede l'altro; si deduce che la caratteristica di tale struttura consiste in un accesso ad un suo elemento tramite una ricerca sequenziale dal primo all'ultimo termine.

La lista può essere di lunghezza fissa, quando il numero dei dati resta costante nel tempo, o variabile e, in tal caso, si parla di struttura dinamica.

Alla lista di lunghezza variabile si associano diversi nomi quali:

- coda;
- pila;
- doppia coda.

Con il termine coda si indica una struttura in cui gli inserimenti avvengono dopo l'ultimo elemento, mentre le estrazioni con eliminazione avvengono sul primo elemento; ciò fa sì che a questa struttura venga attribuito il nome FIFO (**F**irst-**I**n-**F**irst-**O**ut).

Con il termine pila si indica una struttura, detta anche *catasta* o *stack*, in cui l'estrazione e l'introduzione dei dati avvengono dallo stesso estremo; ciò fa sì che a questa struttura venga attribuito in nome LIFO (**L**ast-**I**n-**F**irst-**O**ut).

Con il termine doppia coda si indica una struttura, detta anche *deque*, in cui gli inserimenti e le estrazioni possono avvenire indifferentemente in entrambe gli estremi.

La matrice, detta anche *array*, è un insieme finito di elementi in corrispondenza biunivoca con un insieme di n -ple ordinate di numeri interi detti indici. La matrice è una struttura a lunghezza fissa caratterizzata dall'accesso diretto, all'elemento voluto, tramite la n -pla di indici che lo contraddistingue; nel caso in cui la matrice sia unidimensionale le si dà il nome di vettore.

Con il termine tavola si indica una struttura, detta anche *tabella*, costituita da un insieme finito di elementi, formati da due parti, detti **record**⁽¹⁾. La prima parte è detta chiave, o nome, mentre la seconda è detta valore; ciò fa sì, essendo le chiavi diverse tra loro, che l'accesso ad un elemento della tavola avvenga in modo diretto tramite la chiave stessa.

Col termine grafo si indica una struttura costituita da un insieme di punti, detti nodi o vertici, e collegati tra di loro da segmenti, detti lati o spigoli, in grado di rappresentare una struttura astratta di dati qualora ai nodi si assegnino dei dati e ai lati le relazioni che intercorrono tra i dati. Queste strutture possono essere orientate (spigoli con frecce) o meno e sono caratterizzati, a seconda del collegamento tra i nodi, da:

- cammino: successione di nodi adiacenti;
- cammino semplice, ovvero cammino orientato: cammino comprendente nodi distinti;
- cammino ciclo, ovvero cammino ciclo orientato: cammino che congiunge un nodo con se stesso;
- cammino ciclo, o circuito: cammino ciclico semplice.

In caso in cui è sempre possibile congiungere due nodi presi ad arbitrio, il grafo si dice connesso.

Con il termine albero si intende un insieme finito di elementi, detti nodi, che può essere vuoto e da un insieme di nodi ripartiti in m ($m \geq 0$) sottoinsiemi disgiunti, ciascuno dei quali è a sua volta un albero, detti sotto-alberi della radice. Esso può essere definito come un albero che sia connesso e privo di circuiti, ovvero privo di lati multipli.

I nodi, a seconda del tipo di relazione che si stabilisce tra di essi, si dicono:

- radice o padre: nodo dal quale ne derivano altri;
- figlio: nodo derivato da un padre;
- fratello: nodo appartenente alla stessa radice, o padre;
- foglia: nodo privo di sotto-albero.

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presependo> nella home page della sezione dedicata a Melfi.

In un albero i nodi ne caratterizzano il nome, se ordinati (albero ordinato), e il grado (massimo grado raggiunto dai nodi, per i quali è dato dal massimo numero di figli).

Un particolare tipo di albero è l'albero binario che è, a sua volta, definito come un insieme finito di elementi, vuoto o costituito da un nodo, e da un insieme finito di nodi ripartito in due sottoinsiemi distinti detti sotto-albero sinistro e sotto-albero destro. Una particolare caratteristica di tale struttura è il bilanciamento; un albero binario si dice perfettamente bilanciato se per ogni nodo il numero di nodi appartenenti al sotto-albero di sinistra, e al sotto-albero di destra, si differenziano al più di una unità

STRUTTURE CONCRETE DI MEMORIA

GENERALITÀ

La struttura concreta di memoria è una organizzazione complessa, realizzata tramite programmi, riguardante una parte della memoria principale per far sì che le strutture astratte si adattino alla struttura della memoria stessa (celle). Queste organizzazioni sono dette:

- struttura sequenziale;
- struttura concatenata;
- plesso.

La struttura sequenziale, che è la struttura concreta più semplice, è costituita da un insieme di elementi, che sono porzioni di memoria comprendenti più celle, fisicamente adiacenti nella memoria. In una struttura così definita l'accesso ai dati è definito dalla seguente formula:

$$\text{IND}(x_i) = \text{IND}(x_1) + (i - 1)l$$

ove $\text{IND}(x_i)$ è l'indirizzo dell'elemento che si vuol trovare, $\text{IND}(x_1)$ è l'indirizzo iniziale della struttura ed l è la lunghezza degli elementi della struttura.

La struttura concatenata, detta anche catena, è costituita da un insieme di elementi disposti in modo arbitrario nella memoria purché non sovrapposti. Ogni elemento della catena, è diviso in due parti in cui sono memorizzati il dato della struttura astratta e l'indirizzo, detto anche puntatore, dell'elemento successivo.

Si distinguono due tipi di strutture astratte, esse sono:

- catena semplice;
- catena bidirezionale o catena doppia.

Nella prima tipologia di catena, ogni elemento contiene un solo puntatore che indica l'indirizzo dell'elemento successivo; l'ultimo elemento, per indicare che la catena non ha altri elementi, possiede un puntatore nullo.

Nella seconda tipologia, invece, ogni elemento contiene due puntatori che indicano, rispettivamente, l'indirizzo dell'elemento precedente e l'indirizzo dell'elemento successivo; in tal caso anche l'elemento iniziale possiede un puntatore nullo.

Il plesso è una struttura dinamica, che si compone di un insieme di elementi i quali, come per le strutture concatenate, sono disposti in modo arbitrario nella memoria. Ogni elemento si compone di più parti contenenti rispettivamente:

- il dato associato all'elemento;
- un numero variabile di puntatori associati a più elementi;
- descrizione del formato del plesso.

Nel caso in cui un elemento del plesso non presenta puntatori, ad esso si dà il nome di elemento terminale del plesso.

MEMORIZZAZIONE DELLE STRUTTURE ASTRATTE.

Le liste possono essere memorizzate o mediante strutture sequenziali o mediante struttura a catena dipendente dal fatto che sia, o meno, di lunghezza costante. Per cui in caso di lunghezza costante, e le operazioni cui è sottoposta atte solo per l'accesso ai suoi dati, la memorizzazione più opportuna è quella con struttura sequenziale; mentre nel caso in cui le operazioni sono atte a variare la lunghezza della lista, a causa di eliminazioni o aggiunta di elementi, è più opportuna una struttura concatenata.

La coda è generalmente rappresentata tramite catena circolare con puntatore sull'ultimo elemento che, a sua volta, contiene le informazioni del primo elemento. La pila è memorizzata con

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presepano> nella home page della sezione dedicata a Melfi.

strutture sequenziali con puntatore sull'elemento affiorante. La doppia coda si presta a memorizzazione con strutture sequenziali, se pur con qualche problema, e a strutture a catena.

La memorizzazione di una matrice è soggetta alla struttura della matrice stessa. Se essa è una matrice semplice è più opportuna una struttura sequenziale, mentre se è una matrice bidirezionale, a n righe e m colonne, essa può essere vista come un vettore con $n \times m$ elementi aventi accesso diretto determinato dalle seguenti relazioni:

$IND(A_{ij}) = IND(A_{11}) + (i - 1)n + (j - 1)l$; *memorizzazione per righe*

$IND(A_{ij}) = IND(A_{11}) + (j - 1)m + (i - 1)l$; *memorizzazione per colonne*

dove $IND(A_{ij})$ è l'indirizzo di un generico elemento, $IND(A_{11})$ è l'indirizzo del primo elemento della colonna (ovvero riga), l è la lunghezza dell'elemento, m ed n sono rispettivamente colonne e righe.

Tale sistema non è più idoneo per rappresentare matrici di caratteristiche particolari come: matrice sparsa, matrice simmetrica, matrice triangolare poiché, in questi casi, per determinare la struttura di memorizzazione più idonee occorre tenere presente anche il tipo di operazioni a cui sono soggette queste matrici.

Le tavole sono, in genere, rappresentate con una struttura concreta di tipo sequenziale in cui, per l'uso previsto dalle tavole, vi sono diversi metodi di accesso ai singoli elementi come:

- ricerca sequenziale;
- ricerca binaria;
- accesso calcolato.

Il primo metodo, utilizzato quando gli elementi sono utilizzati in un ordine qualsiasi, effettua un confronto tra la chiave voluta e le n chiavi della struttura in modo sequenziale fin quando, dal confronto, non si ottiene l'uguaglianza tra due chiavi. Con questo metodo di accessi risulta semplice aggiungere altri elementi alla struttura.

La ricerca binaria è effettuata nelle strutture in cui gli elementi siano ordinati secondo le chiavi: $K_1 < K_2 < K_3 < \dots < K_n$, anche in questo caso vi è un confronto tra la chiave voluta e quella della struttura che, nella ricerca binaria, coincide con la chiave dell'elemento centrale $(n - 1)/2$ se n dispari, o $n/2$ se n pari. Il risultato del confronto indica in quale metà continuare la ricerca, che prosegue applicando iterativamente la stessa procedura. L'aggiunta di nuovi elementi comporta un riordinamento delle chiavi con costi proibitivi.

L'accesso regolato si basa sulla definizione di una funzione detta funzione di accesso o di parcheggio. Tra i vari tipi di accesso calcolato si distinguono: l'accesso calcolato di tipo aperto, l'accesso calcolato con traboccamento esterno e l'accesso calcolato con concatenamento interno.

La memorizzazione di grafi ed alberi avviene tramite l'uso di strutture concrete quali strutture concatenate o plessi.

Organizzazione degli archivi su nastro magnetico.

Il nastro magnetico si presta alla memorizzazione di strutture sequenziali aventi accesso di tipo sequenziale. l'archivio memorizzato, che è visto come una lista di record, viene memorizzato per blocchi o per record fisici. Come il nastro può contenere più archivi, ed è detto bobina multiarchivio, un archivio può essere contenuto in un solo nastro (archivio monovolume) o in più nastri (archivio multivolume).

In genere la memorizzazione dei dati su questi supporti può così essere schematizzata:

- **una etichetta di volume:** per l'identificare la bobina, che contiene il numero di matricola della bobina ed il nome del proprietario;
- **una etichetta di archivio:** che contiene informazioni sull'archivio che la segue come
 1. codice e numero identificativo dell'archivio;
 2. numero di matricola dell'etichetta della bobina;
 3. numero d'ordine dell'archivio (se multivolume);
 4. data di registrazione dell'archivio;
 5. data di scadenza dell'archivio.
- **un tape mark:** che permette di posizionare la testina di lettura e scrittura all'inizio dell'archivio;
- **blocchi contenenti i record;**

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presepano> nella home page della sezione dedicata a Melfi.

- un tape mark;
- etichetta di fine volume (*end of volume*) o etichetta di fine archivio (*end of file*);
- uno o due tape mark.

I record possono essere di uguale lunghezza o di lunghezza variabile; in questo caso nei primi caratteri di ciascun record dovrà essere memorizzata l'indicazione della lunghezza del record stesso.

LINGUAGGI SIMBOLICI ORIENTATI ALLA MACCHINA.

TRADUZIONE DI UN PROGRAMMA SORGENTE.

Ogni programma, scritto in un qualsiasi linguaggio simbolico, per essere riconosciuto ed utilizzato da un elaboratore deve essere tradotto nel corrispondente linguaggio macchina. I programmi che svolgono queste operazioni sono detti assemblatori e possono essere:

- assemblatori a una passata;
- assemblatori a due passate.

Gli assemblatori a una passata hanno l'inconveniente di bloccarsi in caso di comandi di salto, o in caso di operandi simbolici non definiti in precedenza. La traduzione del programma sorgente avviene istruzione per istruzione.

Gli assemblatori a due passate sono generalmente usati in quei programmi i cui vi sono dei comandi di salto come **GOTO**. È così detto perché, per effettuare la traduzione del programma, compie due scansioni dette passate.

Le funzioni fondamentali svolte da un assemblatore sono:

- analisi sintattica delle istruzioni simboliche;
- calcolo degli indirizzi;
- traduzione delle istruzioni simboliche dichiarate ed esecutive.

L'esecuzione della prima istruzione richiede la consultazione di tabelle contenenti i codici simbolici del programma sorgente. Il calcolo degli indirizzi relativi di memoria è condotto tramite un calcolatore di indirizzi, detto *location counter*, che viene incrementato ad ogni istruzione sorgente di una quantità uguale alla lunghezza della istruzione oggetto corrispondente.

Dopo l'esecuzione di queste due fasi avviene la traduzione vera e propria del programma sorgente.

SISTEMI OPERATIVI

Il sistema operativo è il gestore delle risorse hardware e software dell'elaboratore elettronico. Per la sua analizzazione si usa, generalmente, il modello Onion Skin (buccia di cipolla), elaborato da H. M. Deitel nel 1983, che lo suddivide in sei strati diversi in cui il livello di astrazione dell'hardware è tanto più elevato tanto più ci si allontana dal nucleo. Questi livelli sono detti:

- gestore del processore;
- gestore della memoria;
- periferiche virtuali;
- file system;
- programmi di libreria;
- interfaccia utente.

Nel primo livello vi sono tutti quei moduli che gestiscono il processore allo scopo di segnalare particolarità hardware, del software che lo utilizzerà, e di realizzare il multitasking.

Il secondo livello è gestito dai moduli che gestiscono la memoria centrale che consentono, tra l'altro, anche a più programmi in esecuzione di condividere la memoria stessa.

Il terzo livello si costituisce dei driver appartenenti alle varie periferiche con cui deve dialogare il calcolatore.

Il quarto livello è costituito dal file system che ha il ruolo di memorizzare permanentemente i dati; la memorizzazione fisica, in una memoria di massa, avviene in blocchi fisici distinti. Un altro importante ruolo del file system consiste nel creare l'illusione di una organizzazione gerarchica dei dati.

Il quinto livello è costituito da utility che hanno il compito di collegare i livelli inferiori con le applicazioni in esecuzione sul sistema digitale.

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presependo> nella home page della sezione dedicata a Melfi.

Il sesto livello consente il dialogo tra l'utente e l'elaboratore elettronico. In passato era detto interprete di comandi e consentiva l'uso dell'elaboratore solo ad utenti specialisti; oggi le interfacce grafiche hanno ampliato l'uso dell'elaboratore anche ad utenti meno esperti.

I sistemi operativi si suddividono in due grandi categorie per cui si distinguono S.O. monotasking e sistemi operativi multitasking.

Un sistema operativo monotasking fa sì che, nell'elaboratore, venga eseguito un solo programma per volta; per tal motivo l'esecuzione di un secondo programma potrà avvenire solo al termine del primo causando una sottoutilizzazione delle risorse del computer. In genere, un programma in esecuzione (detto **task** o **processo**) utilizza le seguenti risorse:

- memoria;
- processore;
- unità di I/O.

ovviamente queste tre risorse sono sottoutilizzate perché, sicuramente, la memoria non è completamente utilizzata dal primo processo, il processore può trovarsi in attesa di input dell'utente e le unità di I/O non sempre sono richieste dai programmi.

Un valido esempio di sistema operativo multitasking è l'MS-DOS.

I sistemi operativi multitasking sono in grado di eseguire più programmi per volta dando l'illusione che ciò avvenga contemporaneamente. Le sue principali caratteristiche sono:

- accesso condiviso alle risorse dell'elaboratore;
- memoria virtuale;
- interfaccia utente evoluta.

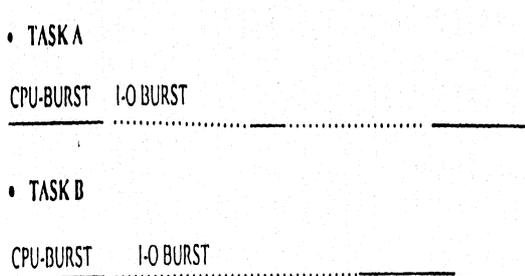
In pratica i sistemi operativi multitasking sono in grado di sfruttare i tempi morti permettendo al processore di svolgere altre operazioni; un valido esempio di sistema operativo multitasking è Windows.

Il task.

Un task è, per definizione, un programma definito da due componenti dette: componente statica, che è costituita dal codice del programma relativo, e componente dinamica che, a sua volta, si compone di:

- uno stato interno: insieme di valori dai registri del processore;
- uno stato di avanzamento: del tipo pronto, attesa o esecuzione;
- priorità.

La priorità può essere di tipo statico, se il suo valore rimane costante durante tutta l'esecuzione del processo, o di tipo dinamico.



Funzionalità dei sistemi multitasking.

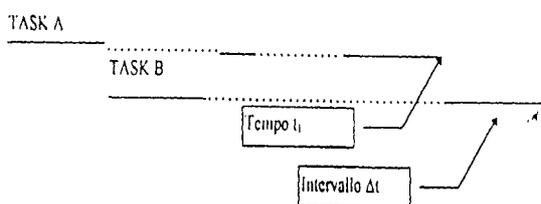
Dalla immagine si nota che il task A esegue le seguenti fasi:

- inizio di elaborazione CPU-Burst;
- fase di I/O;
- elaborazione;
- fase di I/O;
- elaborazione e fine del processo;

mentre la il task B esegue le seguenti fasi:

- inizio elaborazione CPU-Burst;
- fase di I/O;
- elaborazione e fine del processo.

In ambiente monotasking il tempo necessario al completamento del processo, detto **turnaround time (TT)**, è di t_1 per la task A e $t_1 + t_2$ per la task B; mentre in ambiente multitasking il TT è di t_1



per la task A e di $t_1 + \Delta t < t_1 + t_2$ per la task B, dove Δt è il tempo che intercorre tra la fine del processo A e la fine del processo B. Il numero dei processi che un computer è in grado di eseguire in parallelo definisce il livello di multiprogrammazione.

Algoritmi di scheduling.

Il file è di proprietà di quipo.it. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presepano> nella home page della sezione dedicata a Melfi.

Gli algoritmi di scheduling hanno l'obiettivo di minimizzare i tempi di attesa (**wait time WT**) e le medie dei TT. A tal fine il long-time scheduler regola un giusto equilibrio tra task di tipo CPU-Burst e task di tipo I/O-Burst.

Esistono diversi tipi di algoritmo di scheduling, la loro scelta è dettata dal contesto in cui deve operare il sistema operativo; i più noti sono:

- algoritmo FIFO o FCFS (**F**irst-**C**ome-**F**irst-**S**erved);
- algoritmo SJF (**S**hortes-**J**ob-**F**irst);
- algoritmo di scheduling basato sulla priorità del task.

L'algoritmo più semplice è quello denominato FIFO, esso fa sì che il primo processo che arriva allo stato di **ready** sia il primo ad utilizzare il processore.

È comunque possibile che, per una differenza di tempi o per assenza di fasi di I/O tra più task, non sia sempre il primo ad aver usufruito del processore che termini per primo; in questo caso si parla più correttamente di FCFS. Con questo tipo di algoritmo il primo processo che arriva, pur essendo il primo ad essere servito dal processore, non è detto che sia anche il primo a terminare.

L'algoritmo SJF effettua l'esplorazione dei processi **pronti** (della coda) e la selezione del task con il CPU-Burst più breve. Ciò fa in modo da ridurre i tempi di attesa (WT) di tutti i task della coda riducendo al massimo le medie dei wait time e dei turnaround time. L'unica difficoltà, della sua realizzazione, consiste nel determinare l'esatta coda dei **pronti**.

L'algoritmo di scheduling basato sulla priorità della task, fa in modo che i processi vengano eseguiti per valori di priorità decrescenti. Al fine di evitare che i processi con bassi valori di priorità non vengano mai eseguiti (**startvation o morte per fame di CPU**), a tutti i processi vengono incrementati i valori della priorità; in tal modo, prima o poi, anche i task con valori bassi di priorità potranno essere serviti dalla CPU.

Algoritmo di Round Robin.

L'algoritmo di Round Robin, a differenza dei precedenti in cui i task si impossessano del processore, assegna la CPU, a rotazione, ai vari task in coda per un tempo massimo prestabilito detto **time slice**. Ciò a sì che la CPU venga rilasciata dal processo in tre casi:

- durante il time slice in caso di operazioni di I/O (il task va nello stato di **WAIT**);
- durante il time slice se il task termina la sua esecuzione (il task va nello stato di **STOP**);
- quando finisce il time slice (il task va nella coda dei **PRONTI**).

L'efficienza di tale algoritmo è soggetta al valore del time slice; per valori molto ampi il suo funzionamento è simile all'FCFS, mentre per valori troppo piccoli si ha un appesantimento, nonché un degrado, delle prestazioni del sistema.

Rilocazione.

Col termine rilocazione si indica la capacità del sistema di posizionare ed eseguire i processi in qualsiasi parte della memoria centrale.

Il meccanismo di rilocazione fa in modo che gli indirizzi, generati dal programma in esecuzione (**indirizzi logici**), vengano generati in modo corretto. Dopo tale operazione, svolge la traduzione dell'indirizzo logico in indirizzo fisico, per poter trovare la posizione reale dell'indirizzo stesso, tramite la somma tra l'indirizzo generato dal programma e l'indirizzo della primo locazione in cui risiede il programma stesso.

Nucleo del sistema operativo.

Il nucleo del sistema operativo, detto anche monitor, è una parte del codice del sistema operativo stesso che, in fase di avviamento (**start-up**), viene caricato nella parte alta della memoria centrale. Questa zona è vietata all'accesso da parte dei task per cui, al fine di salvaguardare tale zona, il sistema operativo confronta gli indirizzi delle task con quello del **fence register**⁽²⁾ e, in caso di riferimento alla zona del nucleo, viene generato un **errore di indirizzamento in memoria**.

Frammentazione della memoria.

La frammentazione è un indicatore della quantità di memoria centrale inutilizzata a causa della strategia di gestione della memoria adottata e di una particolare configurazione di processi presenti nel sistema.

Vi sono due tipi di frammentazione, essi sono:

- frammentazione esterna;

Il file è di proprietà di Pierpaolo N. Se ne autorizza la copia purché non sia a scopo di lucro. Il file è scaricabile dal mio sito web all'indirizzo: <http://utenti.quipo.it/presepano> nella home page della sezione dedicata a Melfi.

- frammentazione interna.

La frammentazione esterna misura la quantità di memoria di una intera partizione rimasta inutilizzata, mentre la partizione interna misura la quantità di memoria resa inutilizzata all'interno di una partizione.

Le partizioni possono essere di due tipi differenti che sono: partizione fissa e partizione variabile.

La memoria centrale a partizione fissa consiste in divisioni rigide della memoria non modificabili.

I task vengono inseriti tramite tre diversi algoritmi nei seguenti modi:

- assegnazione di un task nella partizione di memoria più ampia della quantità di memoria richiesta dal task stesso;
- assegnazione delle partizioni di memoria secondo le priorità dei task;
- assegnazione del primo processo alla prima partizione in grado di ospitarlo.

Il sistema a partizione fissa non ha problemi di frammentazione esterna ma ha notevoli problemi di frammentazione interna.

La memoria centrale a partizione variabile consiste in una divisione non fissa della memoria, variabile in relazione alla dimensione del processo che deve ospitare.

I task vengono inseriti secondo il livello di priorità in modo decrescente. Con questo sistema non si hanno problemi di frammentazione interna ma si hanno dei problemi di frammentazione esterna.

Riorganizzazione della memoria.

La frammentazione esterna cresce proporzionalmente al numero di task eseguiti rendendo così necessaria la riorganizzazione della memoria al fine di renderla più ordinata. Questo metodo consente di eliminare la frammentazione esterna creando una partizione di dimensione uguale alla somma delle dimensioni dei buchi creati in memoria. Questo processo, gestito da un algoritmo, può essere eseguito nei seguenti modi:

- ad intervalli di tempo costanti;
- quando si crea una frammentazione maggiore o uguale alla frammentazione di soglia;
- quando si carica un task in memoria centrale;
- quando un task rilascia la memoria centrale.

L'efficienza di tale algoritmo è importante poiché consente di ridurre notevolmente il tempo necessario alla riorganizzazione della memoria.

LA MEMORIA VIRTUALE.

La memoria virtuale è un'illusione creata dal sistema operativo. Il suo scopo è quello di caricare parzialmente un programma in memoria, facendo in modo da ridurre al massimo il problema della frammentazione esterna ed eliminare del tutto la riorganizzazione della memoria. Esistono varie tecniche di realizzazione della memoria virtuale, le più comuni sono la **paginazione**⁽³⁾ e la **segmentazione**.

La paginazione consiste in una suddivisione di pagine logiche del programma (di uguali dimensioni) sparse nella memoria che, a sua volta, è divisa in pagine fisiche. La segmentazione si differenzia dalla paginazione per la modalità con cui viene formata la partizione che deve contenere il codice del programma da eseguire.

Queste tecniche permettono di eseguire programmi aventi dimensioni maggiori della dimensione della memoria centrale.