PROGRAMMA ASSEMBLER

È costituito da due sezioni, una sezione dati (.DATA) e una sezione testo (.TEXT)
Un programma assembler può essere memorizzato su più file (file sorgente) ed in ognuno di potranno essere messuna una o più sezioni dati e sezioni testo.

Solo in uno dei file assembler deve esserci l'entry-point dell'intero programma (_start) che deve essere dichiarato globale. Questo è il file principale mentre gli altri sono file secondari.

Nel file principale sono contenuti: il programma principale, le istruzioni di ritorno al sistema operativo ed eventuali sottoprogrammi.
Nei file secondari ci sono solo sottoprogrammi

Il risultato del programma (per il sistema operativo) deve essere lasciato in EBX: se vale 0 tutto si è svolto correttamente altrimenti contiene un valore intero che rappresenta la codifica dell'errore

Le istruzioni di ritorno sono:

```
MOVL ...,%ebx
MOVL $1,%eax    # n° d'ordine della primitiva exit
                # parametro per la INT
INT  $0x80
```

e vanno eseguite alla fine del file che contiene l'entry_point.

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

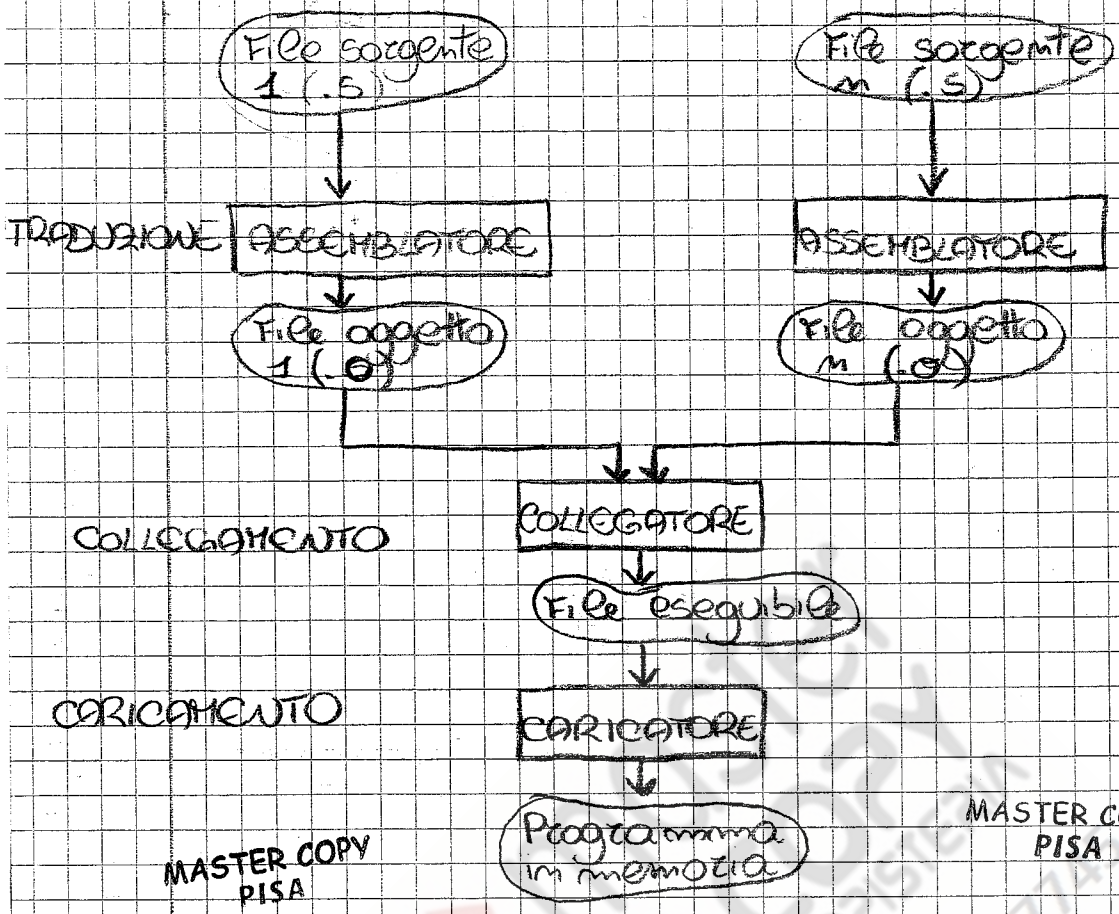
MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

MASTER COPY
PISA

PROGRAMMA ASSEMBLER SU PIÙ FILE

File sorgente: - ha estensione .S
- viene tradotto dando luogo a un file oggetto

Ogni file sorgente viene tradotto dando luogo a un file oggetto ciascuno.
I file oggetto vengono poi collegati per formare un unico eseguibile che verrà caricato in memoria a partire da un indirizzo prestabilito.
Il caricatore inizializza il puntatore di pila (ESP) e il contatore di programma (EIP) (ponendo in pila il valore dell'etichetta _start) ed esegue l'istruzione RET



Vediamo i comandi per le varie fasi:

TRADUZIONE: as --32 nome_file1.s -o nome_file1.o

as --32 nome_filen.s -o nome_filen.o

COLLEGAMENTO: ld -melf_i386 nome_file1.o ... nome_filen.o -o nome_file

CARICAMENTO ED ESECUZIONE: ./nome_file

Oppure con il comando g++

g++ -m32 nome_file1.s ... nome_filen.s -o nome_file

Tale comando richiama l'assemblatore per ognuno dei file elencati e poi il collegatore per ognuno (più eventuali librerie del C++).
Tra questi c'è il file che contiene l'entry-point che richiama il sottoprogramma main e ritorna al sistema operativo.

Se il programma è predisposto per il comando g++ il file principale contiene il nome main.

NOMI ESTERNI E NOMI GLOBALI

NOMI ESTERNI: (.EXTERN) nomi che vengono usati ma sono definiti in altri file

NOMI GLOBALI: (.GLOBAL) nomi che possono essere utilizzati anche in altri file

REGOLE DI COLLEGAMENTO

- In un file i nomi delle variabili e i nomi delle funzioni definite fuori dalle funzioni sono implicitamente globali, possono quindi essere riferiti da altri file.
- In un file si possono riferire nomi definiti in altri file purché siano dichiarati esterni nel file e globali nel file dove sono definiti.

MASTER COPY PISA

PROGRAMMAZIONE MISTA

MASTER COPY PISA

I compilatori C/C++ e l'assemblatore producono dei file oggetto compatibili per cui è possibile scrivere i file sorgenti in uno qualsiasi di questi linguaggi traducendoli poi con il proprio traduttore.
Si utilizza poi un unico collegatore che produce l'eseguibile.

I compilatori C/C++ traducono in assembler:

- Le variabili globali in C/C++ sono variabili della sezione dati globali.
- Le funzioni C/C++ sono sottoprogrammi non annidati della sezione testo.
- I parametri e le variabili locali delle funzioni vengono manipolati con la pila.
- Nomi di variabili globali in C/C++ rimangono uguali in assembler
- Nomi di funzioni in C rimangono gli stessi.
- Nomi di funzioni C++ corrispondono a nomi assembler secondo regole che tengono conto del nome, numero e tipo degli argomenti così da poter avere overloading.
  Se si dichiarano extern si possono usare i nomi C++.

PROGRAMMI DI SVILUPPO

ASSEMBLATORE

Ipotesi:

- Variabili e costanti fanno parte della sezione dati
- Le istruzioni di salto prevedono indirizzamento relativo al valore del contenuto di EIP

Il processo di traduzione avviene in due passate (file per file), per ognuna viene usata una variabile interna di conteggio (CL Contatore di Locazioni) che indica l'indirizzo relativo del comando all'interno della sezione.

All'inizio CL=0. Viene azzerata quando si incontra la direttiva .TEXT e .DATA.
Ad ogni comando viene incrementato del numero di byte richiesti per la traduzione del comando in esame.

Nel file da tradurre le direttive #include vengono risolte prima della trad

PRIMA PASSATA -> esame sequenziale del file da tradurre con avanzamento di CL e costruzione della tabella delle sezioni e della tabella dei simboli

La lunghezza di una sezione viene sempre allineata a multipli di 16 ed è pari al valore di CL dopo l'allineamento.

- TABELLA DELLE SEZIONI

  Nome_file - Sezione Lunghezza

  Viene passata al collegatore

- TABELLA DEI SIMBOLI

  Ind_simbolico Sezione Ind_numer_rel(CL)

  Associa ad ogni nome (ind_simbolico) un indirizzo numerico relativo all'inizio della sezione in cui compare (valore di CL in corrispondenza del comando che ha quel nome.

  Viene usata dall'assemblatore nella seconda passata

MASTER COPY
PISA

SECONDA PASSATA → Traduzione sequenziale del file sorgente e predisposizione di altre tabelle per il collegatore.

Il file tradotto è formato da due sezioni ciascuna composta da una sequenza di byte (ad ognuno è associato un indice che rappresenta l'indirizzo numerico relativo di quel byte)

L'assemblatore calcola le espressioni-indirizzo che compaiono negli operandi: ad ogni nome esterno associa l'indirizzo 0, ad ogni nome definito nel file associa l'indirizzo relativo prelevato dalla tabella dei simboli.

- Traduzione istruzioni operative con espressioni indirizzo:
  - INDIRIZZAMENTO DIRETTO: espressione-indirizzo calcolata, valore posto nel campo DISP
  - INDIRIZZAMENTO IMMEDIATO: espressione calcolata, valore posto nel campo IMM.
- Traduzione istruzioni di controllo con indirizzamento relativo rispetto a EIP
  - espressione indirizzo calcolata (E)
  - DISP = E - IND*
    - DISP può essere lungo 1 byte o 1 long.
    - IND* → indirizzo del comando successivo all'istruzione di controllo (EIP quando l'istruzione viene eseguita)

- TABELLA DEI NOMI GLOBALI
  sottoinsieme della tabella dei simboli; contiene i nomi dichiarati globali nel file.

  Nome file - Sezione Nome Globale (CL)

- TABELLA DEI NOMI ESTERNI

  Nome_file Nome_esterno

- TABELLA DI COLLEGAMENTO specifica la sezione e l'indirizzo numerico relativo delle parole lunghe che vanno modificate in fase di collegamento. Servono modifiche perché l'assemblatore calcola provvisoriamente le espressioni-indirizzo.

MASTER COPY PISA
MASTER COPY PISA