

Il processo di Assembling/Compilazione: Assembler, Compiler, Linker, Loader

Corrado Santoro

Dipartimento di Matematica e Informatica
santoro@dmf.unict.it



Corso di Architettura degli Elaboratori

Sviluppo software (in assembly)

- **Editing**: Scrittura del codice sorgente
- **Assembling**: traduzione del sorgente in codice macchina
- **Linking**: inclusione delle librerie
- **Loading**: caricamento ed esecuzione

Assembler

- L'**Assembler** è un tool software che analizza un testo contenente codice sorgente assembly e produce un file binario contenente le istruzioni in codice macchina
- Il processo consiste in una **interpretazione** (parsing) del testo allo scopo di verificare la corretta sintassi delle istruzioni e generare il **bytecode** equivalente
- Ogni riga viene analizzata **in sequenza** e interpretata
- Il risultato è la produzione di un **file “oggetto”** (object file) che contiene sì il codice macchina, ma esso non è direttamente eseguibile

Risoluzione delle Label

- Una parte importante del processo di assembling è la **risoluzione delle etichette (label)**
- Ogni etichetta deve essere sostituita, nel bytecode, dall'indirizzo equivalente
- Tale indirizzo deve essere calcolato rispetto a dove inizia effettivamente il nostro codice
- L'indirizzo di inizio è noto come **Origin Address** o semplicemente **Origin**
- Esso può essere noto a:
 - **compile/assemble time**
 - **link time**
 - **run time**

Risoluzione delle Label

```
org 0x8000

mov r0, #10
loop:
    sub r0, r0, #1
    cmp r0, #0
    bne loop
...
```

- L'etichetta **loop** è posizionata all'indirizzo **0x8004**, pertanto l'istruzione di branch verrà trasformata come:

bne 0x8004

Risoluzione delle Label

```
    org 0x8000

    mov r0, #10
loop:
    cmp r0, #0
    beq end_loop
    sub r0, r0, #1
    b loop
end_loop:
    ....
```

- L'etichetta **end_loop** non può essere risolta quando l'assemblatore analizza l'istruzione **beq**

Two-Pass Assembly

```
    org 0x8000

    mov r0, #10
loop:
    cmp r0, #0
    beq end_loop
    sub r0, r0, #1
    b loop
end_loop:
    ....
```

- Il processo di assembling prevede in genere **due passi**
- **Passo 1:** Si analizzano le istruzioni individuando la loro occupazione in memoria e si marcano le etichette con il loro indirizzo
- **Passo 2:** Si risolvono tutte le istruzioni che fanno riferimento a etichette in quanto adesso è noto l'indirizzo di ogni etichetta

Assembling e Linking

Multi-Part Assembling

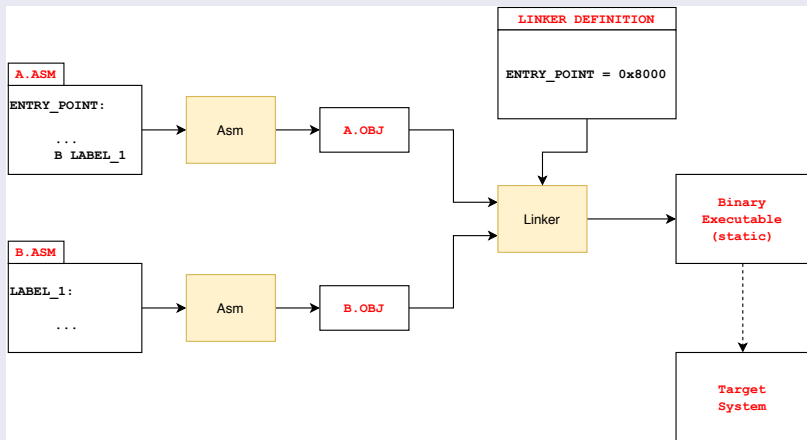
- Un programma in Assembly può essere costituito da più file sorgenti, e non necessariamente da un unico file
- In questi casi è sempre frequente che ci siano dei **riferimenti mutui**: il file **A** può contenere un **branch** ad una etichetta definita nel file **B**
- Poichè ogni file viene assemblato singolarmente, non è possibile in alcun modo (in fase di assembling) risolvere le etichette

Assembling e Linking

- Ogni file viene assemblato singolarmente e si produce un **object file**:
 - $A.asm \Rightarrow A.obj$
 - $B.asm \Rightarrow B.obj$
- Ogni object file contiene anche i **riferimenti** delle singole etichette
- Gli object file vengono poi “fusi insieme” nel processo di **linking**

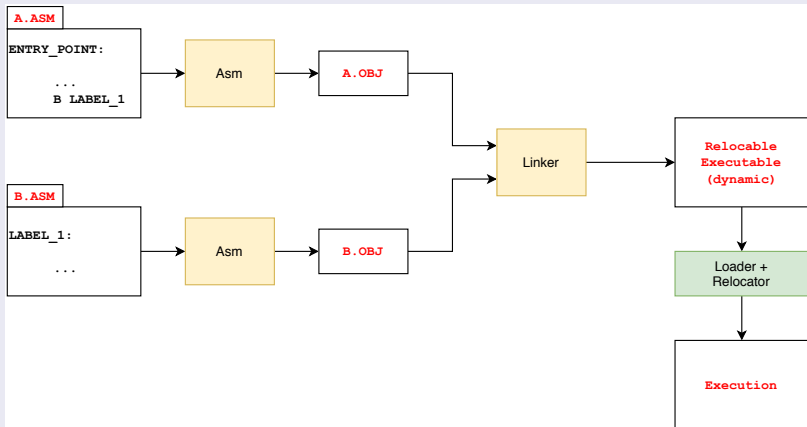
Assembling e Static Linking

Assembling e Static Linking



Assembling e Dynamic Linking

Assembling e Dynamic Linking



Compiler

- Il **compilatore** è un tool software che analizza un testo contenente un codice sorgente di un **linguaggio di alto livello** e produce un file binario contenente le istruzioni in codice macchina
- Il processo consiste in una **interpretazione** (parsing) degli statement del linguaggio e la produzione di un equivalente **sorgente assembly**
- Il sorgente assembly viene dunque analizzato dall'assembler e infine dal linker che produce l'eseguibile
- Negli ambienti di sviluppo abitualmente utilizzati, questo processo è direttamente eseguito avviando il compilatore:

```
$ g++ hello.cpp -o hello
```

Il processo di Assembling/Compilazione: Assembler, Compiler, Linker, Loader

Corrado Santoro

Dipartimento di Matematica e Informatica
santoro@dmf.unict.it



Corso di Architettura degli Elaboratori