

# Compiler/Interprete: Front-End e Back-End

Front-End  
(Compilatore  
Interprete)

SYMBOL TABLE	
1	position    . . .
2	initial    . . .
3	rate      . . .
4	

Back-End  
Compilatore

Aho, A.V. et al., Compilers: Principles,  
Techniques, & Tools, 2 ed., Addison-Wesley,  
2007, pag. 7, Fig. 1.7

position := initial + rate \* 60

lexical analyzer

id<sub>1</sub> := id<sub>2</sub> + id<sub>3</sub> \* 60

syntax analyzer

id<sub>1</sub> :=  
id<sub>2</sub> +  
id<sub>3</sub> \* 60

semantic analyzer

id<sub>1</sub> :=  
id<sub>2</sub> +  
id<sub>3</sub> \* intoreal  
60

intermediate code generator

temp1 := intoreal(60)  
temp2 := id3 \* temp1  
temp3 := id2 + temp2  
id1 := temp3

code optimizer

temp1 := id3 \* 60.0  
id1 := id2 + temp1

code generator

MOVF id3, R2  
MULF #60.0, R2  
MOVF id2, R1  
ADDF R2, R1  
MOVF R1, id1

- **Programmi (sintatticamente) Legali** sono i programmi sintatticamente corretti a cui la semantica del linguaggio può associare la funzione calcolabile descritta dal programma.
- La sintassi espressa da una grammatica non ambigua non è sufficiente a identificare tutti e soli i programmi legali.
- Ad esempio:  $x = 7$  è un assegnamento
  - sintatticamente corretto in tutti i programmi C<sup>1</sup>
  - ma non è legale se  $x$  non è stato *dichiarato*<sup>2</sup>

---

<sup>1</sup>e di tutti i L.P. dove  $x$ ,  $_=_$ ,  $7$  siano la sintassi per una variabile, per l'operatore di assegnamento, per un'espressione

<sup>2</sup>opportunamente, nel programma

- **Proprietà contestuali** (delle regole di composizione) non possono essere espresse da una grammatica libera (da contesto, i.e. Context-Free):
  - identificatori usati devono essere dichiarati prima ...;
  - numero di parametri attuali e formali devono coincidere
  - compatibilità nell'assegnamento, tra il tipo di una variabile e il tipo dell'espressione assegnata
  - ...
- Queste proprietà contestuali **fanno parte della definizione del linguaggio**, ma devono essere espressi con strumenti adatti.
- **Analisi Statica** Sono procedimenti di analisi e formalismi specifici (ad es., sistema dei tipi) per controllare che il programma soddisfi tutti le proprietà contestuali del linguaggio.

- **Programmi (sintatticamente) Legali.** Compilatori e Interpreti trattano tutti questi aspetti nel front-end.
- Il **Front-End** di C./I.<sup>3</sup> provvede alla:
  - **Analisi Lessicale** in accordo al lessico
  - **Analisi Sintattica** in accordo alla sintassi
  - **Analisi Statica** in accordo alle proprietà contestuali
  - **Costruzione Abstract Tree**  $AT(p)$  che fornisce una rappresentazione interna del programma sorgente  $p$ .
- $AT(p)$  è successivamente utilizzato dal **Back-End** di C./I. per completare la realizzazione dell'esecutore.
- Vediamo tutto questo nella tipica struttura (in fasi) di un C.

---

<sup>3</sup>Compilatori e Interpreti

- Associare significato ai termini del linguaggio.
- Vari formalismi con differenti accezioni di significato e differenti usi.
- Due da ricordare:
  - **Semantica Denotazionale**
    - *Significato*: Funzione Calcolata
    - *Usi*: Molteplici incluso definizione di C. e di I.
    - *Laboratorio*: Interpreti di  $\mathcal{L}$ , derivabili dalla S. Den. di  $\mathcal{L}$ , riscritta in un Linguaggio di P. Funzionale, OCaml
    - *Caratteristica*: Orientata allo studio di proprietà di  $\mathcal{L}$ , Astratta dall'implementazione di  $\mathcal{L}$ .
  - **Semantica Operazionale**

- Associare significato ai termini del linguaggio.
- Vari formalismi con differenti accezioni di significato e differenti usi.
- Due da ricordare:
  - **Semantica Denotazionale**
  - **Semantica Operazionale**
    - *Significato*: Computazione (su Macchina)
    - *Usi*: Definizione di I., anche didattici e per studio di programmi e programmazione in  $\mathcal{L}$
    - *Caratteristica*: Basata sulla nozione di Stato, transizione di stato, computazione
- Useremo una Semantica Operazionale *Strutturata* in cui la Macchina è una Macchina Astratta.

- La definizione di Stato, Transizione di stato e Computazione di una S.O.S, dipende dal Linguaggio.
- Vediamole per il Linguaggio sotto.
- Assumiamo che sintassi (concreta) e lessico siano già stati trattati, analizzati, fornendoci la sintassi astratta dei termini del linguaggio: "(AExp-Aexp)" va letto come un albero con radice "-" e due alberi "AExp" come figli (i simboli "(,")" rimarcano questa lettura).

$$Num ::= 1 \mid 2 \mid 3 \mid \dots$$
$$Var ::= X_1 \mid X_2 \mid X_3 \mid \dots$$
$$AExp ::= Num \mid Var \mid (AExp + AExp) \mid (AExp - AExp)$$
$$BExp ::= \mathbf{tt} \mid \mathbf{ff} \mid (AExp == AExp) \mid \neg BExp \mid (BExp \wedge BExp)$$
$$Com ::= \mathbf{skip} \mid Var := AExp \mid Com; Com \mid$$
$$\mathbf{if} BExp \mathbf{then} Com \mathbf{else} Com \mid \mathbf{while} BExp \mathbf{do} Com$$

$$Num ::= 1 \mid 2 \mid 3 \mid \dots$$
$$Var ::= X_1 \mid X_2 \mid X_3 \mid \dots$$
$$AExp ::= Num \mid Var \mid (AExp + AExp) \mid (AExp - AExp)$$
$$BExp ::= \mathbf{tt} \mid \mathbf{ff} \mid (AExp == AExp) \mid \neg BExp \mid (BExp \wedge BExp)$$
$$Com ::= \mathbf{skip} \mid Var := AExp \mid Com; Com \mid$$
$$\mathbf{if} BExp \mathbf{then} Com \mathbf{else} Com \mid \mathbf{while} BExp \mathbf{do} Com$$

- La definizione di Stato.
  - Memoria simbolica per trattare le variabili (valori modificabili)
  - Non abbiamo I/O, file systems,... Lo stato è la sola memoria
  - Lo stato è rappresentato da sequenza finita di coppie  $(X_i, n_i)$ , indicante ...
  - Lo stato è denotato con le variabili  $\sigma, \tau$  (anche con pedici)

$$Num ::= 1 \mid 2 \mid 3 \mid \dots$$
$$Var ::= X_1 \mid X_2 \mid X_3 \mid \dots$$
$$AExp ::= Num \mid Var \mid (AExp + AExp) \mid (AExp - AExp)$$
$$BExp ::= \mathbf{tt} \mid \mathbf{ff} \mid (AExp == AExp) \mid \neg BExp \mid (BExp \wedge BExp)$$
$$Com ::= \mathbf{skip} \mid Var := AExp \mid Com; Com \mid$$
$$\mathbf{if} BExp \mathbf{then} Com \mathbf{else} Com \mid \mathbf{while} BExp \mathbf{do} Com$$

- La definizione di Transizione.
  - Esecuzione di un costrutto  $c$  nello stato  $\sigma$  della Macchina Astratta
  - La esprimiamo con:
    - $\langle c, \sigma \rangle \rightarrow \tau$ , indicante ...
    - $\langle c, \sigma \rangle \rightarrow \langle c', \sigma' \rangle$ , indicante ...
    - $\langle c_1, \sigma_1 \rangle \rightarrow \langle c'_1, \sigma'_1 \rangle, \dots, \langle c_k, \sigma_k \rangle \rightarrow \langle c'_k, \sigma'_k \rangle / \langle c, \sigma \rangle \rightarrow \langle c', \sigma' \rangle$ ,  
k-premesse/1-conclusione, indicante ...
  - La semantica di  $\mathcal{L}$  fornisce le transizioni che sono specifiche per  $\mathcal{L}$

$$Num ::= 1 \mid 2 \mid 3 \mid \dots$$
$$Var ::= X_1 \mid X_2 \mid X_3 \mid \dots$$
$$AExp ::= Num \mid Var \mid (AExp + AExp) \mid (AExp - AExp)$$
$$BExp ::= \mathbf{tt} \mid \mathbf{ff} \mid (AExp == AExp) \mid \neg BExp \mid (BExp \wedge BExp)$$
$$Com ::= \mathbf{skip} \mid Var := AExp \mid Com; Com \mid$$
$$\mathbf{if } BExp \mathbf{ then } Com \mathbf{ else } Com \mid \mathbf{while } BExp \mathbf{ do } Com$$

- La definizione di Computazione di  $p \in \mathcal{L}$ .
  - Sequenza degli stati attraversati effettivamente dalle transizioni usate nell'esecuzione del programma  $p$ .

$$Num ::= 1 \mid 2 \mid 3 \mid \dots$$
$$Var ::= X_1 \mid X_2 \mid X_3 \mid \dots$$
$$AExp ::= Num \mid Var \mid (AExp + AExp) \mid (AExp - AExp)$$
$$BExp ::= \mathbf{tt} \mid \mathbf{ff} \mid (AExp == AExp) \mid \neg BExp \mid (BExp \wedge BExp)$$
$$Com ::= \mathbf{skip} \mid Var := AExp \mid Com; Com \mid$$
$$\mathbf{if} BExp \mathbf{then} Com \mathbf{else} Com \mid \mathbf{while} BExp \mathbf{do} Com$$

- Notazione.

$(X_1, n_1), \dots, (X_k, n_k)$  stato con  $k$  variabili legate

$\sigma, \tau$  (anche con pedici) sono stati della macchina

$\sigma(X_i) = n_i$  quando  $\sigma = (X_1, n_1), \dots, (X_i, n_i), \dots, (X_k, n_k)$ ,

$\sigma[X_i \leftarrow m_i] = (X_1, n_1), \dots, (X_i, m_i), \dots, (X_k, n_k)$

quando  $\sigma = (X_1, n_1), \dots, (X_i, n_i), \dots, (X_k, n_k)$

$n, n_i \in Num$  valori numerici

$a, a_i \in AExp$  espressioni aritmetiche

$b, b_i \in BExp$  espressioni booleane

$\mathbf{tt}, \mathbf{ff}$  i valori true e false

$c, c_i \in Com$  comandi del linguaggio

$$\text{Num} ::= 1 \mid 2 \mid 3 \mid \dots$$
$$\text{Var} ::= X_1 \mid X_2 \mid X_3 \mid \dots$$
$$\text{AExp} ::= \text{Num} \mid \text{Var} \mid (\text{AExp} + \text{AExp}) \mid (\text{AExp} - \text{AExp})$$

## Semantica delle Espressioni AExp.

$$\langle X, \sigma \rangle \rightarrow \langle \sigma(X), \sigma \rangle$$
$$\frac{\langle n + m, \sigma \rangle \rightarrow \langle p, \sigma \rangle}{\text{where } p = n + m}$$
$$\frac{\langle n - m, \sigma \rangle \rightarrow \langle p, \sigma \rangle}{\text{where } p = n - m \text{ e } n \geq m}$$
$$\frac{\langle a_1, \sigma \rangle \rightarrow \langle a', \sigma \rangle}{\langle (a_1 + a_2), \sigma \rangle \rightarrow \langle (a' + a_2), \sigma \rangle} \quad \frac{\langle a_2, \sigma \rangle \rightarrow \langle a'', \sigma \rangle}{\langle (a_1 + a_2), \sigma \rangle \rightarrow \langle (a_1 + a''), \sigma \rangle}$$
$$\frac{\langle a_1, \sigma \rangle \rightarrow \langle a', \sigma \rangle}{\langle (a_1 - a_2), \sigma \rangle \rightarrow \langle (a' - a_2), \sigma \rangle} \quad \frac{\langle a_2, \sigma \rangle \rightarrow \langle a'', \sigma \rangle}{\langle (a_1 - a_2), \sigma \rangle \rightarrow \langle (a_1 - a''), \sigma \rangle}$$

- L'ordine di valutazione degli argomenti delle operazioni è inessenziale
- Se lo volessimo da sinistra a destra: Come modificarla?

$$\text{Com} ::= \text{skip} \mid \text{Var} := A\text{Exp} \mid \text{Com}; \text{Com} \mid$$
$$\text{if } B\text{Exp} \text{ then } \text{Com} \text{ else } \text{Com} \mid \text{while } B\text{Exp} \text{ do } \text{Com}$$

## Semantica dei Comandi Com.

$$\langle \text{skip}, \sigma \rangle \rightarrow \sigma \quad (c1)$$

$$\langle X := n, \sigma \rangle \rightarrow \sigma[X \leftarrow n] \quad (c2) \quad \frac{\langle a, \sigma \rangle \rightarrow \langle a', \sigma \rangle}{\langle X := a, \sigma \rangle \rightarrow \langle X := a', \sigma \rangle} \quad (c3)$$

$$\frac{\langle c_1, \sigma \rangle \rightarrow \sigma'}{\langle c_1; c_2, \sigma \rangle \rightarrow \langle c_2, \sigma' \rangle} \quad (c4) \quad \frac{\langle c_1, \sigma \rangle \rightarrow \langle c'_1, \sigma' \rangle}{\langle c_1; c_2, \sigma \rangle \rightarrow \langle c'_1; c_2, \sigma' \rangle} \quad (c5)$$

$$\langle \text{if tt then } c_1 \text{ else } c_2, \sigma \rangle \rightarrow \langle c_1, \sigma \rangle \quad (c6)$$

$$\langle \text{if ff then } c_1 \text{ else } c_2, \sigma \rangle \rightarrow \langle c_2, \sigma \rangle \quad (c7)$$

$$\frac{\langle b, \sigma \rangle \rightarrow \langle b', \sigma \rangle}{\langle \text{if } b \text{ then } c_1 \text{ else } c_2, \sigma \rangle \rightarrow \langle \text{if } b' \text{ then } c_1 \text{ else } c_2, \sigma \rangle} \quad (c8)$$

$$\langle \text{while } b \text{ do } c, \sigma \rangle \rightarrow \langle \text{if } b \text{ then } c; \text{while } b \text{ do } c \text{ else skip}, \sigma \rangle \quad (c9)$$

# SOS: Semantica dei Comandi

$$\langle \text{skip}, \sigma \rangle \rightarrow \sigma \quad (c1)$$

$$\langle X := n, \sigma \rangle \rightarrow \sigma[X \leftarrow n] \quad (c2) \quad \frac{\langle a, \sigma \rangle \rightarrow \langle a', \sigma \rangle}{\langle X := a, \sigma \rangle \rightarrow \langle X := a', \sigma \rangle} \quad (c3)$$

$$\frac{\langle c_1, \sigma \rangle \rightarrow \sigma'}{\langle c_1; c_2, \sigma \rangle \rightarrow \langle c_2, \sigma' \rangle} \quad (c4) \quad \frac{\langle c_1, \sigma \rangle \rightarrow \langle c'_1, \sigma' \rangle}{\langle c_1; c_2, \sigma \rangle \rightarrow \langle c'_1; c_2, \sigma' \rangle} \quad (c5)$$

$$\langle \text{if tt then } c_1 \text{ else } c_2, \sigma \rangle \rightarrow \langle c_1, \sigma \rangle \quad (c6)$$

$$\langle \text{if ff then } c_1 \text{ else } c_2, \sigma \rangle \rightarrow \langle c_2, \sigma \rangle \quad (c7)$$

$$\frac{\langle b, \sigma \rangle \rightarrow \langle b', \sigma \rangle}{\langle \text{if } b \text{ then } c_1 \text{ else } c_2, \sigma \rangle \rightarrow \langle \text{if } b' \text{ then } c_1 \text{ else } c_2, \sigma \rangle} \quad (c8)$$

$$\langle \text{while } b \text{ do } c, \sigma \rangle \rightarrow \langle \text{if } b \text{ then } c; \text{ while } b \text{ do } c \text{ else skip}, \sigma \rangle \quad (c9)$$

Sia  $c$  la sequenza di comandi:  $X := 1; \text{while } \neg(X == 0) \text{ do } X := X - 1$

$$\begin{aligned} & \langle c, \sigma \rangle \\ & \rightarrow \langle c', \sigma[X \leftarrow 1] \rangle \\ & \rightarrow \langle \text{if } \neg(X == 0) \text{ then } X := (X - 1); c' \text{ else skip}, \sigma[X \leftarrow 1] \rangle \\ & \rightarrow \langle \text{if } \neg(1 == 0) \text{ then } X := (X - 1); c' \text{ else skip}, \sigma[X \leftarrow 1] \rangle \\ & \rightarrow \langle \text{if } \neg\text{ff then } X := (X - 1); c' \text{ else skip}, \sigma[X \leftarrow 1] \rangle \\ & \rightarrow \langle \text{if tt then } X := (X - 1); c' \text{ else skip}, \sigma[X \leftarrow 1] \rangle \\ & \rightarrow \langle X := (X - 1); c', \sigma[X \leftarrow 1] \rangle \\ & \rightarrow \langle X := (1 - 1); c', \sigma[X \leftarrow 1] \rangle \\ & \rightarrow \langle X := 0; c', \sigma[X \leftarrow 1] \rangle \\ & \rightarrow \langle c', \sigma[X \leftarrow 0] \rangle \\ & \rightarrow \langle \text{if } \neg(X == 0) \text{ then } X := (X - 1); c' \text{ else skip}, \sigma[X \leftarrow 0] \rangle \\ & \rightarrow \langle \text{if } \neg(0 == 0) \text{ then } X := (X - 1); c' \text{ else skip}, \sigma[X \leftarrow 0] \rangle \\ & \rightarrow \langle \text{if } \neg\text{tt then } X := (X - 1); c' \text{ else skip}, \sigma[X \leftarrow 0] \rangle \\ & \rightarrow \langle \text{if ff then } X := (X - 1); c' \text{ else skip}, \sigma[X \leftarrow 0] \rangle \\ & \rightarrow \langle \text{skip}, \sigma[X \leftarrow 0] \rangle \\ & \rightarrow \sigma[X \leftarrow 0] \end{aligned}$$

- Esercizio 2.10.1

- (a) Si completi la definizione di una grammatica le cui produzioni sono sotto.
- (b) Si mostri che la grammatica ottenuta è ambigua.

$$\begin{aligned} E &\rightarrow T \mid T + E \mid T - E \\ T &\rightarrow A \mid A * E \end{aligned}$$

Soluzione

(a)  $G = \dots$

(b) Consideriamo la stringa:  $s=A*A+A$ . Per essa possiamo mostrare i due diversi parse tree sotto aventi come frontiera  $s$

- Esercizio 2.10.4

Si dia una grammatica non ambigua che generi tutte e sole le sequenze di parentesi angolate bilanciate

Soluzione

$$\begin{aligned} S &\rightarrow B S \mid \epsilon \\ B &\rightarrow \langle S \rangle \end{aligned}$$

- Esercizio 2.10.5

Si chiama lineare una grammatica le cui produzioni hanno la forma  $A \rightarrow t B$  oppure  $A \rightarrow t$ , dove  $A, B$  siano non terminali e  $t$  sia un terminale o  $\epsilon$ . Un linguaggio che può essere espresso con una grammatica lineare si chiama *regolare* e può essere riconosciuto da un automa a stati finiti. Si mostri che il linguaggio  $L = \{a^n b^m \mid n, m \geq 1\}$  è regolare.

Soluzione

$$\begin{array}{l} S \rightarrow a S \mid B \\ B \rightarrow b B \mid b \end{array}$$

- Esercizio 2.10.5bis

Si dia una grammatica non ambigua per che il linguaggio  $L = \{a^n b^m \mid n \leq m\}$ .

Soluzione

$$S \rightarrow a S b \mid B$$

$$B \rightarrow b B \mid \epsilon$$

od anche:

$$S \rightarrow A B$$

$$A \rightarrow a A b \mid \epsilon$$

$$B \rightarrow b B \mid \epsilon$$

$$\langle X, \sigma \rangle \rightarrow \langle \sigma(X), \sigma \rangle$$

$$\begin{array}{l} \langle (n + m), \sigma \rangle \rightarrow \langle p, \sigma \rangle \\ \text{where } p = n + m \end{array}$$

$$\begin{array}{l} \langle (n - m), \sigma \rangle \rightarrow \langle p, \sigma \rangle \\ \text{where } p = n - m \text{ e } n \geq m \end{array}$$

$$\frac{\langle a_1, \sigma \rangle \rightarrow \langle a', \sigma \rangle}{\langle (a_1 + a_2), \sigma \rangle \rightarrow \langle (a' + a_2), \sigma \rangle} \quad \frac{\langle a_2, \sigma \rangle \rightarrow \langle a'', \sigma \rangle}{\langle (a_1 + a_2), \sigma \rangle \rightarrow \langle (a_1 + a''), \sigma \rangle}$$

$$\frac{\langle a_1, \sigma \rangle \rightarrow \langle a', \sigma \rangle}{\langle (a_1 - a_2), \sigma \rangle \rightarrow \langle (a' - a_2), \sigma \rangle} \quad \frac{\langle a_2, \sigma \rangle \rightarrow \langle a'', \sigma \rangle}{\langle (a_1 - a_2), \sigma \rangle \rightarrow \langle (a_1 - a''), \sigma \rangle}$$

- Esercizio 2.10.6bis

Si modifichino le regole delle espressioni aritmetiche in modo da prescrivere una valutazione degli argomenti da sinistra a destra per la somma e una da destra a sinistra per la sottrazione.

**Soluzione**

$$\langle \text{skip}, \sigma \rangle \rightarrow \sigma \quad (c1)$$

$$\langle X := n, \sigma \rangle \rightarrow \sigma[X \leftarrow n] \quad (c2) \quad \frac{\langle a, \sigma \rangle \rightarrow \langle a', \sigma \rangle}{\langle X := a, \sigma \rangle \rightarrow \langle X := a', \sigma \rangle} \quad (c3)$$

$$\frac{\langle c_1, \sigma \rangle \rightarrow \sigma'}{\langle c_1; c_2, \sigma \rangle \rightarrow \langle c_2, \sigma' \rangle} \quad (c4) \quad \frac{\langle c_1, \sigma \rangle \rightarrow \langle c'_1, \sigma' \rangle}{\langle c_1; c_2, \sigma \rangle \rightarrow \langle c'_1; c_2, \sigma' \rangle} \quad (c5)$$

$$\langle \text{if tt then } c_1 \text{ else } c_2, \sigma \rangle \rightarrow \langle c_1, \sigma \rangle \quad (c6)$$

$$\langle \text{if ff then } c_1 \text{ else } c_2, \sigma \rangle \rightarrow \langle c_2, \sigma \rangle \quad (c7)$$

$$\frac{\langle b, \sigma \rangle \rightarrow \langle b', \sigma \rangle}{\langle \text{if } b \text{ then } c_1 \text{ else } c_2, \sigma \rangle \rightarrow \langle \text{if } b' \text{ then } c_1 \text{ else } c_2, \sigma \rangle} \quad (c8)$$

$$\langle \text{while } b \text{ do } c, \sigma \rangle \rightarrow \langle \text{if } b \text{ then } c; \text{ while } b \text{ do } c \text{ else skip}, \sigma \rangle \quad (c9)$$

## ● Esercizio 2.10.7bis

Siano  $c$  e  $d$  i seguenti comandi:

$c$ :  $X:=1$ ;  $d$

$d$ : **while**( $X==1$ )**do skip**

Si dia la sequenza di transizioni ottenute a partire dallo stato  $\sigma = (X, 0)$

**Soluzione**