

Universita' degli Studi "G. D'Annunzio"
Chieti - Pescara
Facolta' di Economia - CLEI



**Strumenti Formali per
l'Analisi Lessicale - Sintattica dei
Compilatori**

**Laureanda
Caterina Mandolini**

**Relatore
Stefano Bistarelli**

Anno Accademico 2003-2004

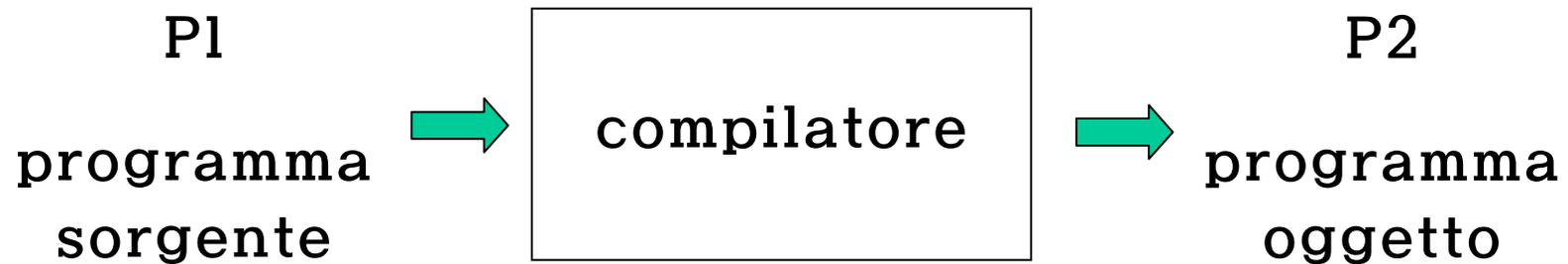


Introduzione

- 1) Compilatori**
- 2) Analisi Lessicale - Sintattica**
- 3) Strumenti Formali**



Il Compilatore come Traduttore



P1 equivalente a P2



Linguaggi di Programmazione

Alto livello: Pascal, C, Java

$X=Y+5$

Basso livello: linguaggio macchina

001 0110
010 0001



Modello analisi - sintesi



ANALISI

Analisi Lessicale
Analisi Sintattica
Analisi Semantica



Il modulo dell'analisi

Vogliamo assegnare alla variabile C il numero 2.795. Supponiamo che l'istruzione corretta sia

`C := 2.795`

`C=2.795` errore!!! (AnLex)

`2.795:=C` errore!!! (Parser)

`C:=2.795` errore!!! (AnSem)



Linguaggi Formali

Il linguaggio matematico e' un linguaggio formale o informale?

$$\forall x \exists y (f(x) < y)$$

La lingua italiana e' un linguaggio formale o informale?

“Conta il numero dei bit”



Linguaggi Formali

E' un insieme costituito da un numero (finito o infinito) di frasi.

Ogni frase e' una sequenza finita di simboli.

$$H = \{f_1, f_2, \dots, f_n, \dots\}$$
$$f_1 = \heartsuit, f_2 = \heartsuit\heartsuit, \dots, f_n = \heartsuit^n, \dots$$

H e' un linguaggio formale?

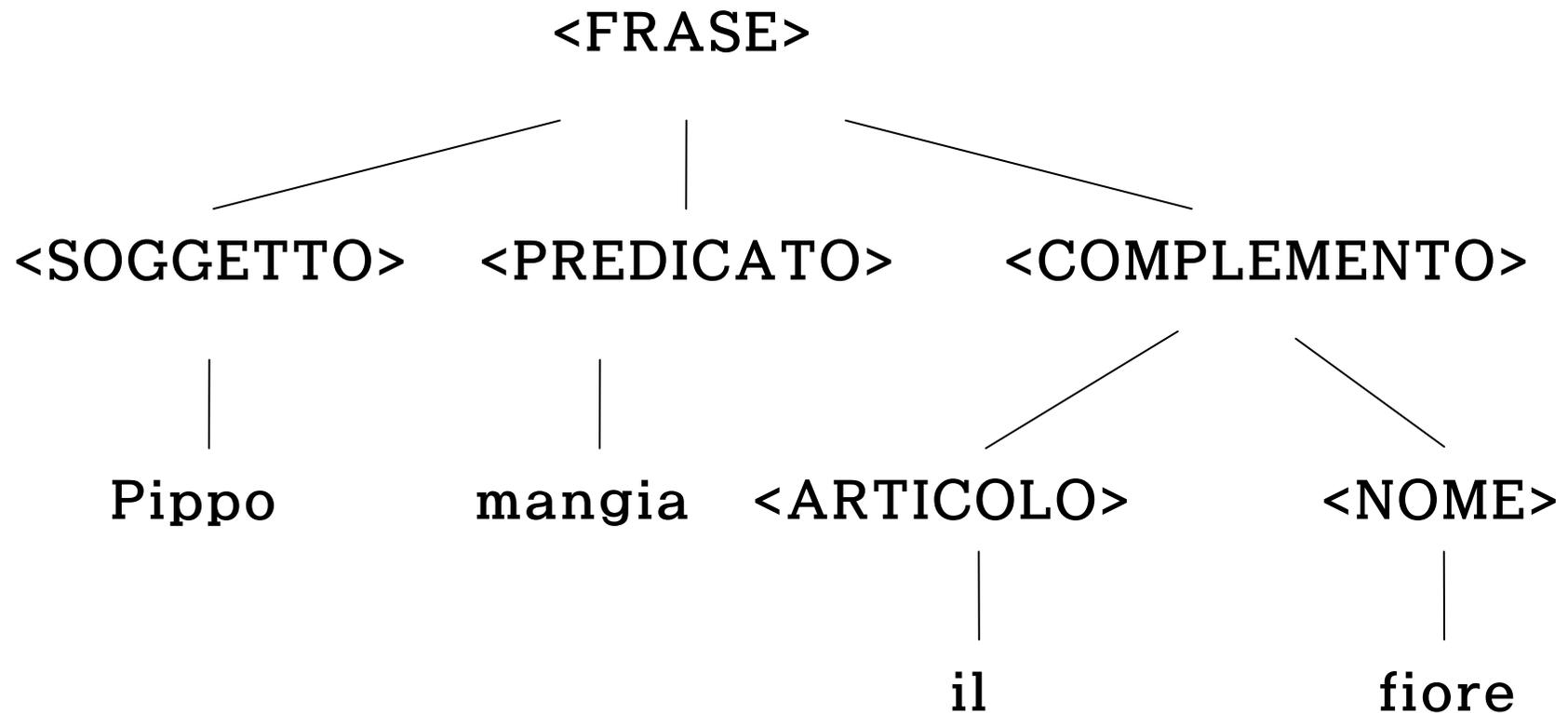


Grammatiche Generative

- Parole
- Categorie sintattiche
- Regole di riscrittura
- Categoria sintattica iniziale



Albero di Derivazione





Classificazione delle Grammatiche alla Chomsky

Tipo 0

$\alpha \rightarrow \beta$

Tipo 1 - Dipendenti dal Contesto

$\gamma A \delta \rightarrow \gamma \beta \delta$

Tipo 2 - Libere dal Contesto

$A \rightarrow \beta$

Tipo 3 - Regolari

$A \rightarrow aB \mid a$



Automa a Stati Finiti

Nastro di input

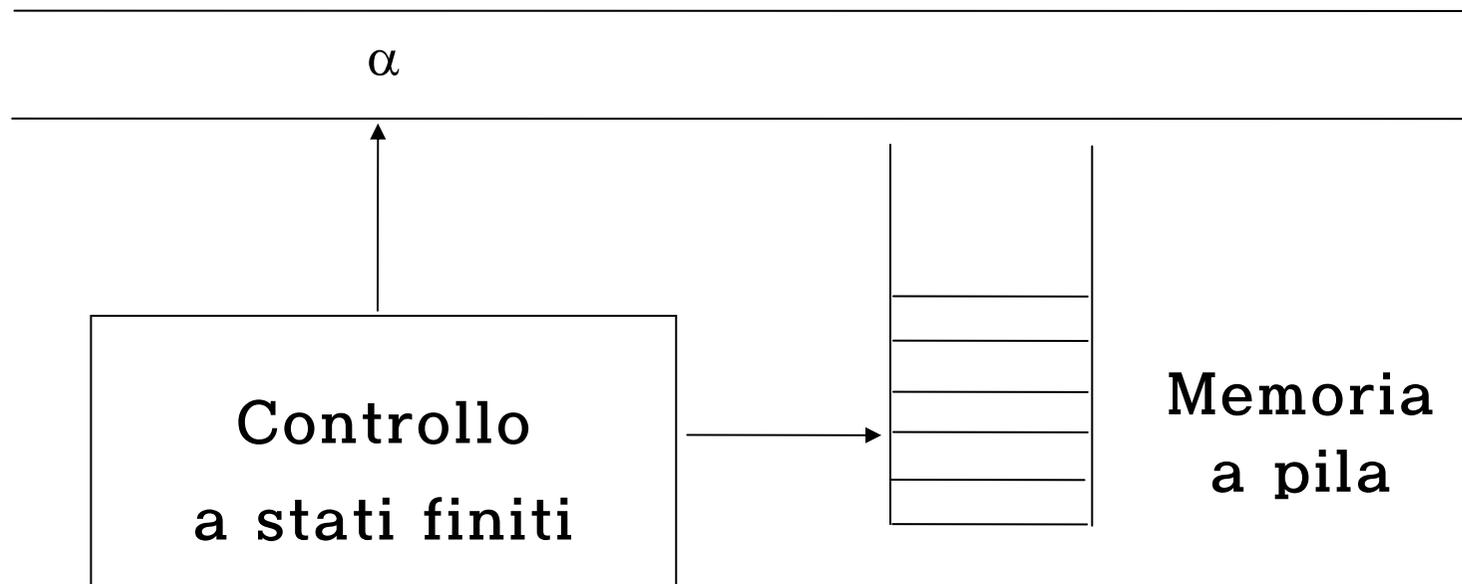
α

Controllo
a stati finiti



Automa a Pila

Nastro di input





Strumenti Automatici

LEX

YACC

(Yet Another Compiler Compiler)



Interazione LEX - YACC





Un Minicompile

Programma YACC

```
% token  NUMERO
```

```
%%
```

```
e: e '+' e  
{ $$ = $1 + $3; }
```

```
e: '(' e ')'  
{ $$ = $2; }
```

```
e: NUMERO  
{ $$ = yylval; }
```

Programma LEX

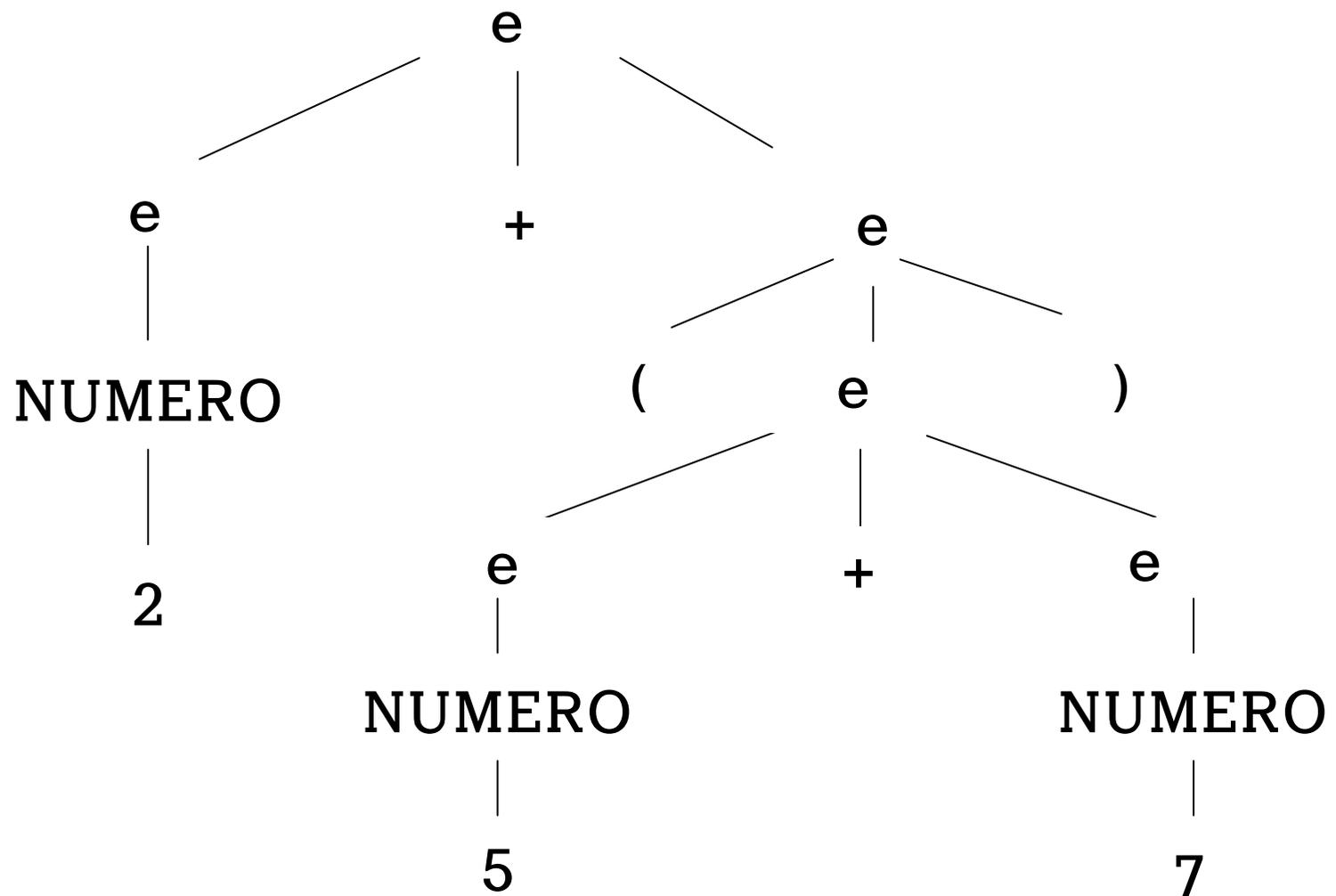
```
NUMERO [0-9]
```

```
%%
```

```
{ NUMERO }  
{ yylval = valore();  
return (NUMERO); }
```



2+(5+7) diviene 14





Conclusioni

**Un programma C ha tante parentesi { quante }.
Un programma Pascal ha tanti BEGIN quanti END.
Da che cosa deriva questa "somiglianza"?**

**Deriva dal fatto che
sono entrambi linguaggi liberi.**



Conclusioni

**Perche' i linguaggi di programmazione
sono linguaggi liberi?**

**Perche' i linguaggi liberi costituiscono
un buon compromesso
tra potere espressivo ed efficienza.**



Conclusioni

Da che cosa deriva la caratteristica dell'efficienza?

Deriva dal fatto che il riconoscimento dei linguaggi liberi puo' essere effettuato tramite l'automa riconoscitore.



Conclusioni

**Quanto costa, in termini di risorse,
implementare l'automa riconoscitore?**

Il costo e' basso utilizzando i tool.



Conclusioni

Realizzare un compilatore
e' un gioco da ragazzi?

ASSOLUTAMENTE NO!



Ringraziamenti

