

Κεφάλαιο 2: **Τυπικές γλώσσες**

(μέρος 2ο)

Γλώσσες χωρίς συμφραζόμενα (i)

- ▶ Γραμματικές χωρίς συμφραζόμενα: $A \rightarrow \alpha$
 - ▶ Σε κάθε παραγωγή **ένα μη τερματικό** σύμβολο αντικαθίσταται, βάσει ενός κανόνα
 - ▶ Πολλές διαφορετικές παραγωγές διαφέρουν μόνο στη **σειρά** των αντικαταστάσεων
- ▶ **Αριστερότερη / δεξιότερη** παραγωγή (leftmost / rightmost derivation)
- ▶ **Συντακτικά δέντρα** (parse trees)

Γλώσσες χωρίς συμφραζόμενα (ii)

- ▶ **Μία** παραγωγή

$$\begin{aligned} [S] &\Rightarrow aA[B]c \Rightarrow a[A]bBc \\ &\Rightarrow acS[B]bBc \Rightarrow ac[S]abBc \\ &\Rightarrow acab[B]c \Rightarrow acabac \end{aligned}$$

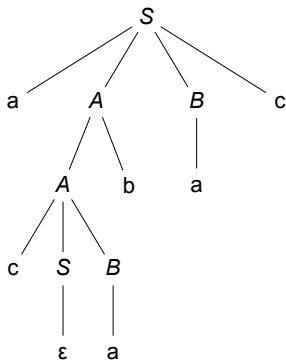
- ▶ **Αριστερότερη** παραγωγή

$$\begin{aligned} [S] &\Rightarrow_L a[A]BC \Rightarrow_L a[A]bBc \\ &\Rightarrow_L ac[S]BbBc \Rightarrow_L ac[B]bBc \\ &\Rightarrow_L acab[B]c \Rightarrow_L acabac \end{aligned}$$

- ▶ **Δεξιότερη** παραγωγή

$$\begin{aligned} [S] &\Rightarrow_R aA[B]c \Rightarrow_R a[A]ac \\ &\Rightarrow_R a[A]bac \Rightarrow_R acS[B]bac \\ &\Rightarrow_R ac[S]abac \Rightarrow_R acabac \end{aligned}$$

$$\begin{aligned} S &\rightarrow aABc \mid \epsilon \\ A &\rightarrow cSB \mid Ab \\ B &\rightarrow bB \mid a \end{aligned}$$



Διφορούμενες γραμματικές (i)

- ▶ Δύο γραμματικές είναι **ισοδύναμες** όταν παράγουν την ίδια γλώσσα.
- ▶ Μια γραμματική είναι **διφορούμενη** (ambiguous) αν υπάρχουν δύο ή περισσότερα συντακτικά δέντρα για την ίδια παραγόμενη συμβολοσειρά
- ▶ Γραμματικές και γλώσσες **εγγενώς διφορούμενες** (inherently ambiguous)
- ▶ Χρήση διφορούμενων γραμματικών στην περιγραφή της σύνταξης γλωσσών προγραμματισμού

Διφορούμενες γραμματικές (ii)

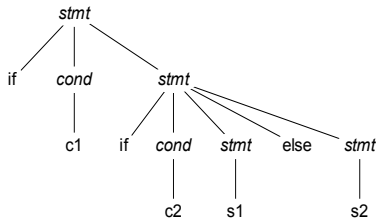
- ▶ Παράδειγμα: **ξηκρέμαστο if** (dangling if)

$stmt \rightarrow \text{if } cond \text{ } stmt \text{ else } stmt \mid \text{if } cond \text{ } stmt \mid s1 \mid s2$

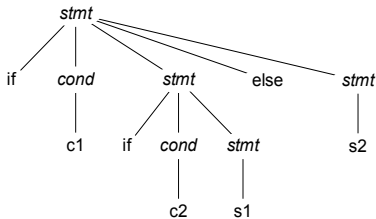
$cond \rightarrow c1 \mid c2$

- ▶ **Διφορούμενο**: σε ποιο if αντιστοιχεί το else;

if c1 if c2 s1 else s2



if c1 (if c2 s1 else s2)



if c1 (if c2 s1) else s2

Τρόποι παράστασης γραμματικών (i)

▶ Backus-Naur Form (BNF)

- ▶ Σύμβολο ::= στους κανόνες
- ▶ Μη τερματικά σύμβολα σε γωνιακές παρενθέσεις, π.χ. ⟨expr⟩
- ▶ Σύμβολο | για διάζευξη

$\langle \text{unsigned-number} \rangle ::= \langle \text{integer-part} \rangle \langle \text{dec-fraction} \rangle \langle \text{exp-part} \rangle$

$\langle \text{integer-part} \rangle ::= \langle \text{digit} \rangle \langle \text{integer-part} \rangle \mid \langle \text{digit} \rangle$

$\langle \text{dec-fraction} \rangle ::= . \langle \text{integer-part} \rangle \mid \epsilon$

$\langle \text{exp-part} \rangle ::= \mathbf{E} \langle \text{sign} \rangle \langle \text{integer-part} \rangle$
 $\mid \mathbf{e} \langle \text{sign} \rangle \langle \text{integer-part} \rangle \mid \epsilon$

$\langle \text{sign} \rangle ::= + \mid - \mid \epsilon$

$\langle \text{digit} \rangle ::= \mathbf{0} \mid \mathbf{1} \mid \mathbf{2} \mid \mathbf{3} \mid \mathbf{4} \mid \mathbf{5} \mid \mathbf{6} \mid \mathbf{7} \mid \mathbf{8} \mid \mathbf{9}$

Τρόποι παράστασης γραμματικών (ii)

▶ Extended Backus-Naur Form (EBNF)

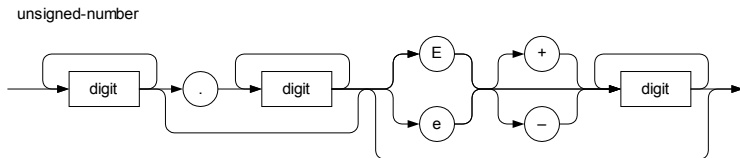
- ▶ Τερματικά σύμβολα σε **εισαγωγικά**
- ▶ **Παρενθέσεις** για ομαδοποίηση
- ▶ **Αγκύλες** για προαιρετικά τμήματα
- ▶ Σύμβολα ***** και **+** για επανάληψη

$\langle \text{unsigned-number} \rangle ::= \langle \text{digit} \rangle^+ [\text{"."} \langle \text{digit} \rangle^+]$
 $[(\text{"E"} | \text{"e"}) [\text{"+"} | \text{"-"}] \langle \text{digit} \rangle^+]$

$\langle \text{digit} \rangle ::= \text{"0"} | \text{"1"} | \text{"2"} | \text{"3"} | \text{"4"}$
 $| \text{"5"} | \text{"6"} | \text{"7"} | \text{"8"} | \text{"9"}$

Τρόποι παράστασης γραμματικών (iii)

- ▶ Συντακτικά διαγράμματα
 - ▶ Τερματικά σύμβολα σε οβάλ
 - ▶ Μη τερματικά σύμβολα σε ορθογώνια
 - ▶ Διαδοχή συμβόλων (παράθεση) με βέλη



Κεφάλαιο 4:

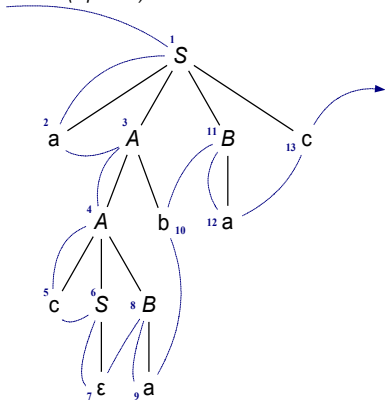
Συντακτική ανάλυση

Συντακτική ανάλυση

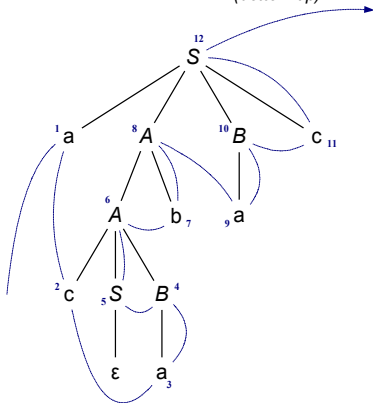
- ▶ **Συντακτικό δέντρο** (parse tree)
- ▶ Κατασκευάζεται με δύο τρόπους:
 - ▶ **Καθοδικά** — Από πάνω προς τα κάτω (**top-down**)
δηλαδή ξεκινώντας **από τη ρίζα** και
προχωρώντας **προς τα φύλλα**
 - ▶ **Ανοδικά** — Από κάτω προς τα πάνω (**bottom-up**)
δηλαδή ξεκινώντας **από τα φύλλα** και
προχωρώντας **προς τη ρίζα**

Top-down και bottom-up

από πάνω προς τα κάτω
(top-down)



από κάτω προς τα πάνω
(bottom-up)



ΣΑ *bottom-up* (i)

- ▶ Η συντακτική ανάλυση ξεκινά από τα φύλλα
- ▶ Κάθε φορά, αναζητά:
 - ▶ τον αριστερότερο κόμβο του δέντρου
 - ▶ που δεν έχει ακόμα κατασκευαστεί
 - ▶ ενώ όλα τα παιδιά του έχουν κατασκευαστεί
- ▶ Επαναλαμβάνει μέχρι να κατασκευαστεί η ρίζα
- ▶ **Ελάττωση** (reducing): η επιλογή των κόμβων που θα αποτελέσουν τα παιδιά ενός νέου κόμβου

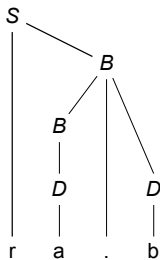
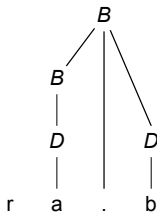
ΣA bottom-up (ii)



$S \rightarrow r B$

$B \rightarrow D \mid B, D$

$D \rightarrow a \mid b$



ΣΑ *bottom-up* (iii)

- ▶ ΣΑ ολίσθησης-ελάττωσης (shift-reduce)
 - ▶ Χρησιμοποιούν μια (αρχικά κενή) **στοίβα** όπου τοποθετούν σύμβολα της γραμματικής
 - ▶ **Ολίσθηση** (shift): μεταφορά ενός συμβόλου από την είσοδο στην κορυφή της στοίβας
 - ▶ **Ελάττωση** (reduce): αφαίρεση από την κορυφή της στοίβας του δεξιού μέλους ενός κανόνα και πρόσθεση του αριστερού μέλους
 - ▶ **Επιτυχία**: η στοίβα περιέχει μόνο το S και τα σύμβολα της εισόδου έχουν εξαντληθεί

ΣΑ bottom-up

(iv)

$S \rightarrow r B$

$B \rightarrow D \mid B , D$

$D \rightarrow a \mid b$

βήμα	στοίβα	είσοδος	κίνηση
0	ϵ	$r a , b$	ολίσθηση
1	r	a , b	ολίσθηση
2	$r a$	$, b$	ελάττωση με $D \rightarrow a$
3	$r D$	$, b$	ελάττωση με $B \rightarrow D$
4	$r B$	$, b$	ολίσθηση (όχι ελάττωση με $S \rightarrow r B$)
5	$r B ,$	b	ολίσθηση
6	$r B , b$	ϵ	ελάττωση με $D \rightarrow b$
7	$r B , D$	ϵ	ελάττωση με $B \rightarrow B , D$ (όχι ελάττωση με $B \rightarrow D$)
8	$r B$	ϵ	ελάττωση με $S \rightarrow r B$
9	S	ϵ	αναγνώριση

ΣΑ *bottom-up* (v)

- ▶ LR(k)
- ▶ LR(0)
- ▶ SLR(1)
- ▶ LALR(1)
- ▶ LR(1)

