

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

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)

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)

$$S \rightarrow aABc \mid \epsilon$$

$$A \rightarrow cSB \mid Ab$$

$$B \rightarrow bB \mid a$$

Γλώσσες χωρίς συμφραζόμενα (ii)

► Μία παραγωγή

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

$$S \rightarrow aABc \mid \epsilon$$

$$A \rightarrow cSB \mid Ab$$

$$B \rightarrow bB \mid a$$

Γλώσσες χωρίς συμφραζόμενα (ii)

▶ Μία παραγωγή

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

▶ Αριστερότερη παραγωγή

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

$$S \rightarrow aABc \mid \epsilon$$

$$A \rightarrow cSB \mid Ab$$

$$B \rightarrow bB \mid a$$

Γλώσσες χωρίς συμφραζόμενα (ii)

▶ Μία παραγωγή

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

▶ Αριστερότερη παραγωγή

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

▶ Δεξιότερη παραγωγή

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

$$S \rightarrow aABc \mid \epsilon$$

$$A \rightarrow cSB \mid Ab$$

$$B \rightarrow bB \mid a$$

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

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

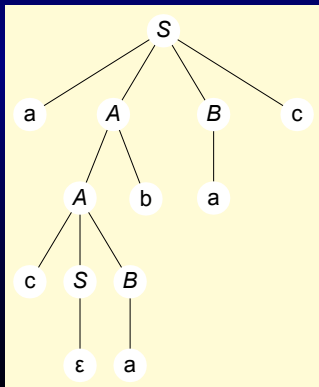
► Δεξιότερη παραγωγή

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

$$S \rightarrow aABc \mid \epsilon$$

$$A \rightarrow cSB \mid Ab$$

$$B \rightarrow bB \mid a$$



Διφορούμενες γραμματικές (i)

- ▶ Δύο γραμματικές είναι **ισοδύναμες** όταν παράγουν την ίδια γλώσσα.

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

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)

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: **ξεκρέμαστο if** (dangling if)

$stmt \rightarrow if\ cond\ stmt\ else\ stmt \mid if\ cond\ stmt \mid s1 \mid s2$

$cond \rightarrow c1 \mid c2$

Διφορούμενες γραμματικές (ii)

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

$stmt \rightarrow if\ cond\ stmt\ else\ stmt \mid if\ cond\ stmt \mid s1 \mid s2$

$cond \rightarrow c1 \mid c2$

- ▶ **Διφορούμενο**: σε ποιο `if` αντιστοιχεί το `else`;
`if c1 if c2 s1 else s2`

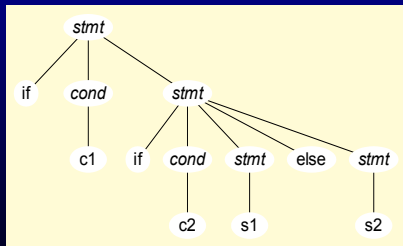
Διφορούμενες γραμματικές (ii)

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

$stmt \rightarrow if\ cond\ stmt\ else\ stmt \mid if\ cond\ stmt \mid s1 \mid s2$

$cond \rightarrow c1 \mid c2$

- ▶ **Διφορούμενο**: σε ποιο if αντιστοιχεί το else;
 $if\ c1\ if\ c2\ s1\ else\ s2$



$if\ c1\ (\ if\ c2\ s1\ else\ s2 \)$

Διφορούμενες γραμματικές (ii)

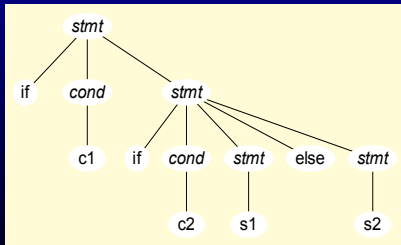
- ▶ Παράδειγμα: **ξεκρέμαστο if** (dangling if)

$stmt \rightarrow if\ cond\ stmt\ else\ stmt \mid if\ cond\ stmt \mid s1 \mid s2$

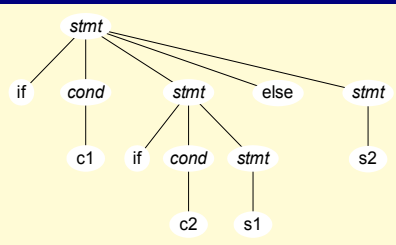
$cond \rightarrow c1 \mid c2$

- ▶ **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

Τρόποι παράστασης γραμματικών

(i)

▶ Backus-Naur Form

(BNF)

- ▶ Σύμβολο ::= στους κανόνες
- ▶ Μη τερματικά σύμβολα σε γωνιακές παρενθέσεις, π.χ. $\langle \text{expr} \rangle$
- ▶ Σύμβολο | για διάζευξη

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

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

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

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

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

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

▶ Extended Backus-Naur Form

(EBNF)

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

▶ Extended Backus-Naur Form

(EBNF)

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

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

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

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

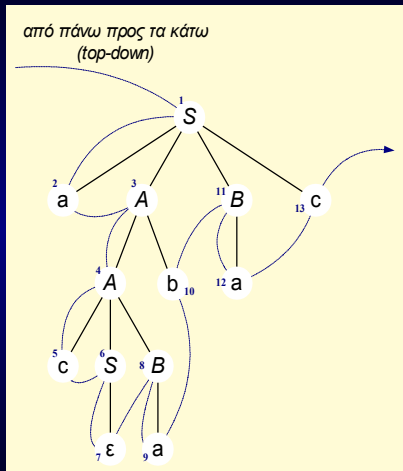
Κεφάλαιο 4:

Συντακτική ανάλυση

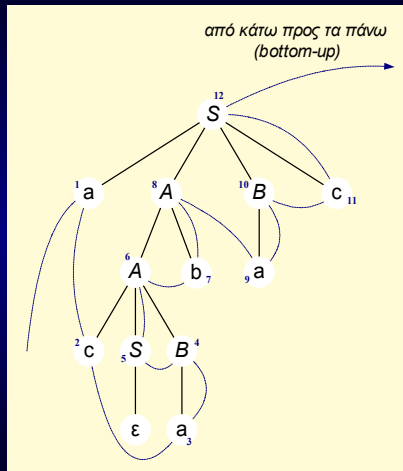
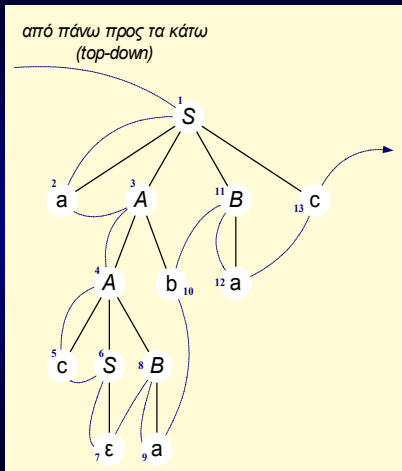
Συντακτική ανάλυση

- ▶ 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



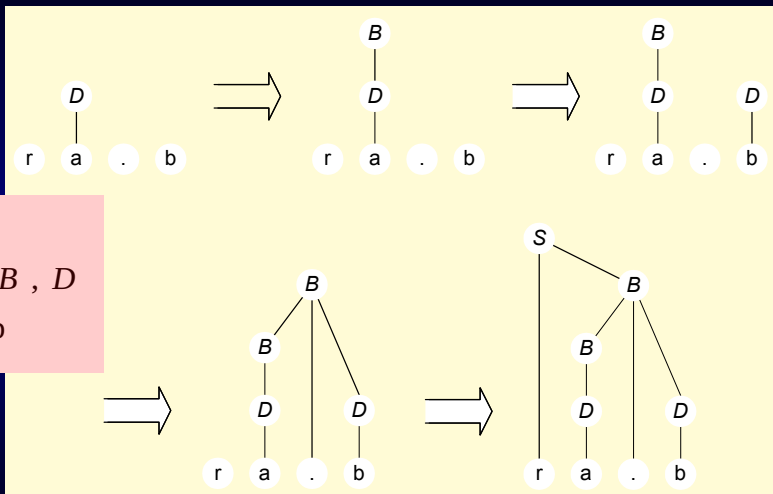
Top-down και bottom-up



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)



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)

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

$$S \rightarrow r B$$

$$B \rightarrow D \mid B, D$$

$$D \rightarrow a \mid b$$

ΣΑ *bottom-up* (v)

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

