

Αλληλεπίδραση Ανθρώπου–Υπολογιστή

B3. Συντακτική και σημασιολογική ανάλυση φυσικής γλώσσας με γραμματικές

(2024-25)

Ίων Ανδρουτσόπουλος

<http://www.aueb.gr/users/ion/>

Οι διαφάνειες αυτής της διάλεξης βασίζονται εν μέρει σε ύλη του βιβλίου «Speech and Language Processing» των D. Jurafsky and J.H. Martin, 2^η έκδοση, Prentice Hall, 2009.

Τι θα ακούσετε

- Γραμματικές χωρίς συμφραζόμενα.
- Κανονική μορφή Chomsky και αλγόριθμος CKY.
- Επαυξημένες γραμματικές χωρίς συμφραζόμενα και γραμματικές DCG της Prolog.
- Σημασιολογική ανάλυση φυσικής γλώσσας με γραμματικές DCG.

Συντακτική ανάλυση με γραμματικές

$s \rightarrow np \ vp$

$np \rightarrow det(\text{Case}) \ n(\text{Case})$

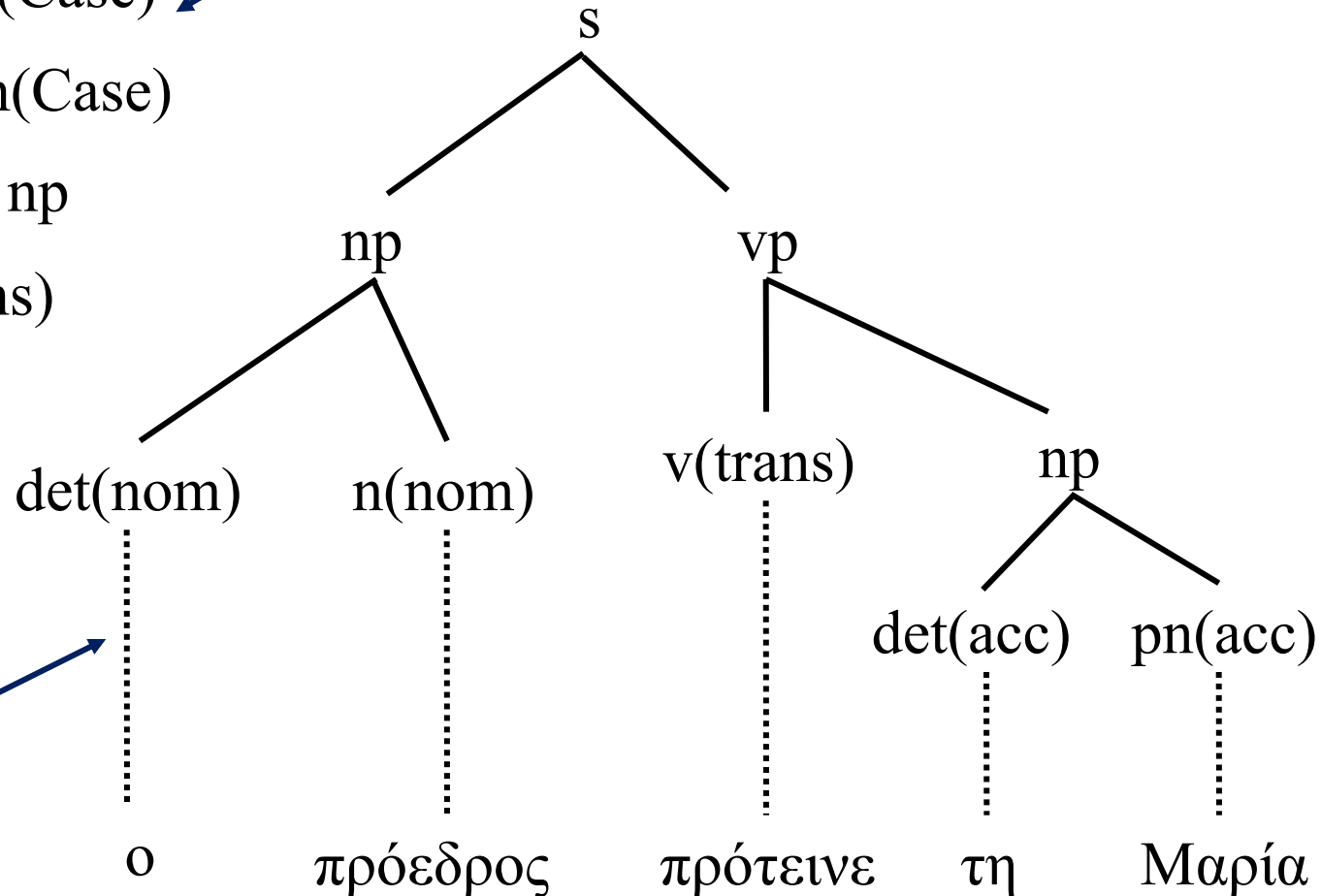
$np \rightarrow det(\text{Case}) \ pn(\text{Case})$

$vp \rightarrow v(\text{trans}) \ np$

$vp \rightarrow v(\text{intrans})$

γραμματική

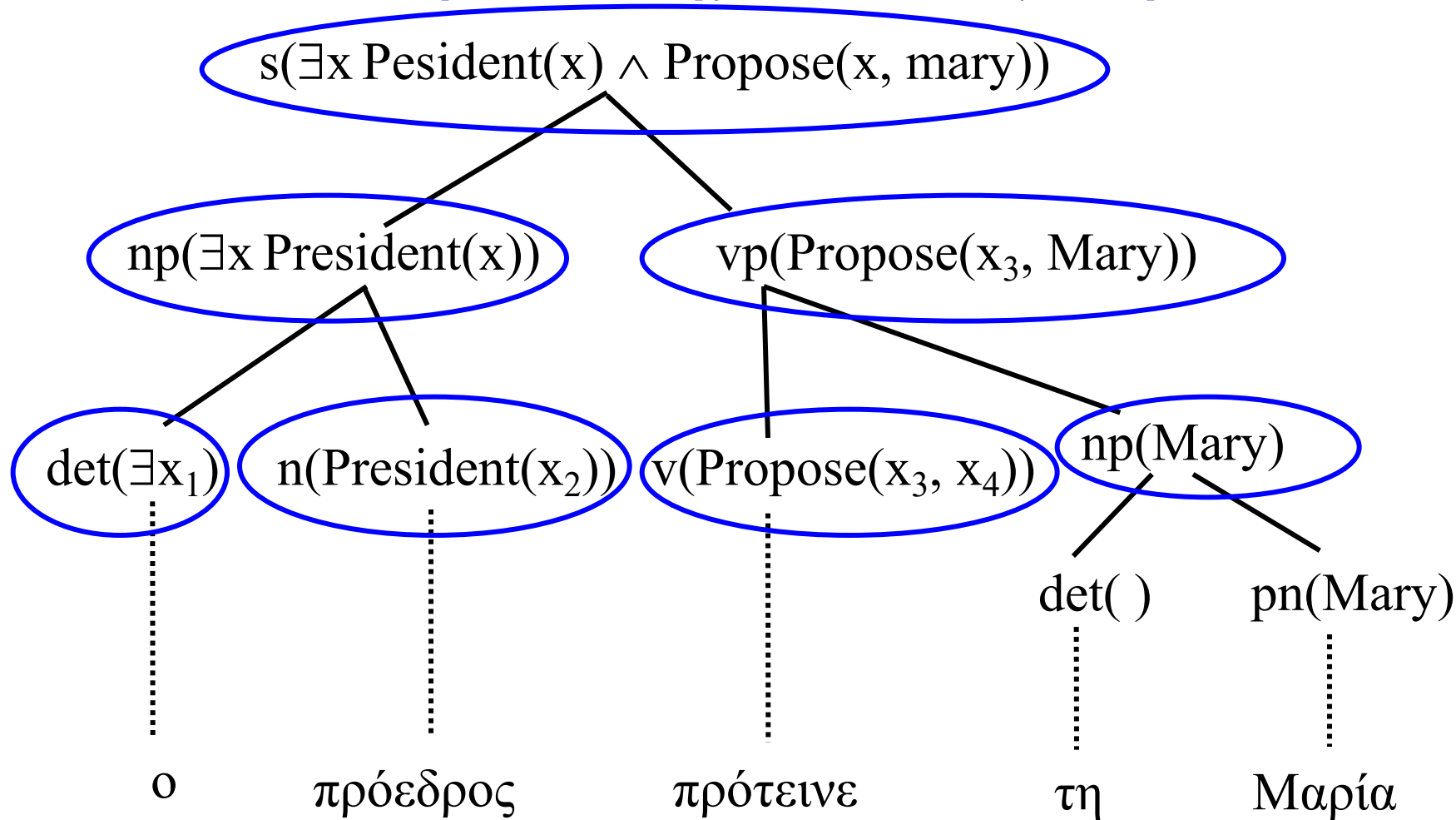
πληροφορίες
μορφολογικής/
λεξιλογικής
ανάλυσης



Σημασιολογική ανάλυση

- Απεικόνιση κάθε πρότασης σε μια παράσταση της **σημασίας** της (του νοήματός της).
 - Π.χ. απεικόνιση σε **λογικές εκφράσεις**.
 - «Ο πρόεδρος πρότεινε τη Μαρία.»
 $\exists x (\text{President}(x) \wedge \text{Propose}(x, \text{Mary}))$
- Παραδοσιακά με προσθήκη τμημάτων χειρισμού **σημασιολογίας** στους κανόνες της γραμματικής.
 - Οι **κανόνες** της γραμματικής επαυξάνονται, ώστε να υπολογίζουν και τη **σημασία** των **κόμβων** του συντακτικού δένδρου.
 - Η **μορφολογική/λεξιλογική** ανάλυση αναλαμβάνει να παρέχει και πληροφορίες για τη **σημασία** των **λέξεων**.

Συντακτική και σημασιολογική ανάλυση



- Θα δούμε στη συνέχεια πιο λεπτομερώς πώς μπορεί να γίνει η συντακτική και σημασιολογική ανάλυση με γραμματικές.

Γραμματικές χωρίς συμφραζόμενα

NP $\rightarrow$ Det Nominal

Nominal $\rightarrow$ N | Adj Nominal

Det $\rightarrow$ ο | η | το | ...

Adj $\rightarrow$ πράσινο | μεγάλο | βαρύ | ...

N $\rightarrow$ βιβλίο | αυτοκίνητο | ...

Διάξευξη. Ουσιαστικά δύο κανόνες.

Μπορούμε να τους
σκεφτούμε ως λεξικό.

- **Τερματικά** σύμβολα, π.χ. «βιβλίο», «το»..
- **Μη τερματικά** σύμβολα, π.χ. «Nominal», «Adj».
- Οι ΓΧΣ έχουν κανόνες της μορφής: $A \rightarrow \alpha$.
 - A : μεμονωμένο **μη τερματικό** σύμβολο.
 - α : (πιθανώς κενή) ακολουθία **τερματικών** και **μη τερματικών**.
- **Αρχικό** σύμβολο: ένα από τα μη τερματικά (εδώ «NP»).
- **Γλώσσα** της γραμματικής: οι **ακολουθίες τερματικών συμβόλων** που παράγονται από το **αρχικό** σύμβολο.

Αλγόριθμοι συντακτικής ανάλυσης

- **Είσοδοι:**

- Μια **γραμματική** του τύπου που υποστηρίζει ο αλγόριθμος (π.χ. γραμματική χωρίς συμφραζόμενα).
- Μια **ακολουθία σ από τερματικά** σύμβολα της γραμματικής.

- **Αποκρίσεις:**

- **Ανήκει** η σ **στη γλώσσα** που ορίζει η γραμματική;
- Ποιο είναι το **συντακτικό δέντρο** της σ ;
- Το συντακτικό δέντρο αποτελεί μια **απόδειξη** ότι η σ είναι σύμφωνη με τη γραμματική. Παρέχει, επίσης, πληροφορίες για τη **συντακτική δομή** της σ .

Μεγαλύτερο παράδειγμα ΓΧΣ

- NP $\rightarrow$ Det PN | Pron | Det Nominal
- Nominal $\rightarrow$ N | Adj Nominal | Nominal PP
- PP $\rightarrow$ Prep NP

- S $\rightarrow$ NP VP | VP
- VP $\rightarrow$ V | V NP

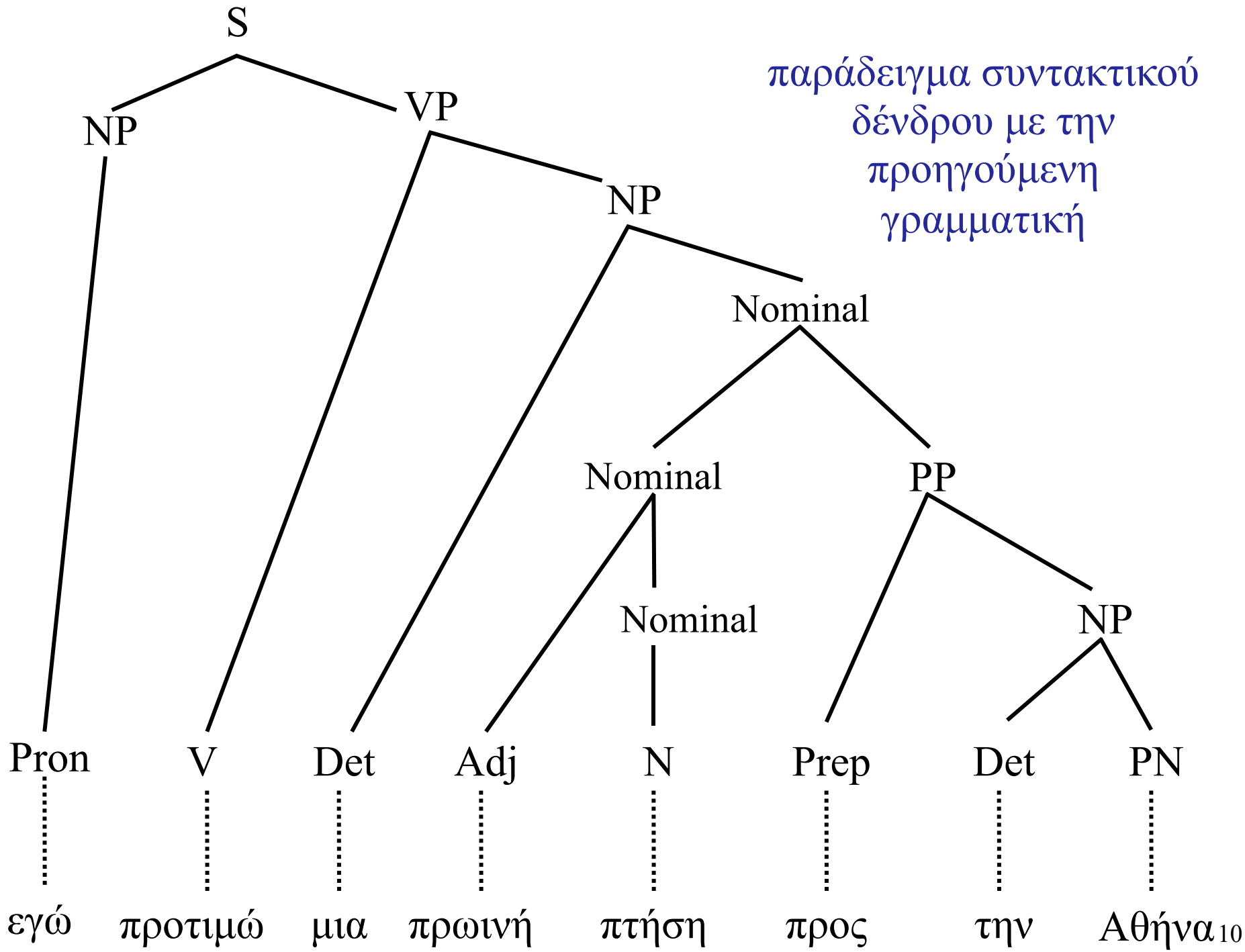
*εγώ θέλω μια
πρωινή πτήση, ...*

*την Αθήνα,
εγώ, μια
πτήση, μια
πρωινή πτήση,
μια πρωινή
πτήση προς την
Αθήνα, ...*

- Pron $\rightarrow$ εγώ
- Det $\rightarrow$ ο | η | έναν | μια | τον | την
- PN $\rightarrow$ Θεσσαλονίκη | Αθήνα
- N $\rightarrow$ πτήση | πελάτης | πελάτη
- Adj $\rightarrow$ πρωινή | απογευματινή
- V $\rightarrow$ θέλω | θέλει | προτιμώ | συμφωνώ
- Prep $\rightarrow$ προς | από

*Παίζουν το ρόλο
λεξικού.*

παράδειγμα συντακτικού
δένδρου με την
προηγούμενη
γραμματική



Συντακτικά διαφορούμενες προτάσεις

- «Είδαμε τον **επιστήμονα με το τηλεσκόπιο.**»
 - Είδαμε [_{NP} τον [_{Nominal} επιστήμονα [_{PP} με το τηλεσκόπιο]]].
 - Όπως «την πτήση από τη Θεσσαλονίκη».
- «**Είδαμε** τον επιστήμονα με το τηλεσκόπιο.»
 - Είδαμε [_{NP} τον επιστήμονα] [_{PP} με το τηλεσκόπιο].
 - Θα είχαμε και κανόνα: $VP \rightarrow V NP PP$.
- «Είδαμε τον επιστήμονα με το τηλεσκόπιο από το Παρίσι.»
 - Είδαμε [τον επιστήμονα] [με το τηλεσκόπιο] [από το Παρίσι].
 - Είδαμε [τον επιστήμονα με το τηλεσκόπιο] [από το Παρίσι].
 - Είδαμε [τον επιστήμονα] [με το [τηλεσκόπιο από το Παρίσι]].
 - Είδαμε [τον [επιστήμονα με το [τηλεσκόπιο από το Παρίσι]]].

Συντακτικά διαφορούμενες προτάσεις

- «Είδαμε τον επιστήμονα **με την άσπρη μπλούζα.**»
 - Χρειαζόμαστε σημασιολογικούς περιορισμούς που να **αποκλείουν** την περίπτωση η **μπλούζα** να είναι το **μέσο** της παρατήρησης.
- Από **καθαρά συντακτική σκοπιά**, οι περισσότερες προτάσεις είναι **εξαιρετικά διαφορούμενες**.
 - Πολύ μεγάλος αριθμός συντακτικών δένδρων (συχνά **εκθετική αύξηση** όσο αυξάνει ο αριθμός των φράσεων που συνδυάζονται).
 - **Χρονοβόρο** να ανακαλύψουμε και να επιστρέψουμε όλα τα συντακτικά δέντρα **ξεχωριστά**.
 - Υπάρχουν **αποδοτικοί αλγόριθμοι** (π.χ. CKY, βλ. παρακάτω) που ουσιαστικά διερευνούν **όλα τα συντακτικά δέντρα ταυτόχρονα** μέσω **δυναμικού προγραμματισμού**.

Κανονική μορφή Chomsky

- Γραμματικές χωρίς συμφραζόμενα σε κανονική μορφή Chomsky (CNF).
 - Επιτρέπονται μόνο κανόνες της μορφής $A \rightarrow B C$ και $A \rightarrow w$, όπου A, B, C μη τερματικά και w τερματικό. Π.χ:

$S \rightarrow V NP$		
$V \rightarrow \text{θέλω}$	$V \rightarrow \text{επιθυμώ}$	
$NP \rightarrow \text{Det Nominal}$	$\text{Nominal} \rightarrow \text{Adj Nominal}$	$\text{Nominal} \rightarrow N$
$\text{Det} \rightarrow \text{μια}$	$N \rightarrow \text{πτήση}$	$\text{Nominal} \rightarrow \text{πτήση}$
$\text{Adj} \rightarrow \text{πρωινή}$	$\text{Adj} \rightarrow \text{απογευματινή}$	
- Κάθε ΓΧΣ μπορεί να μετατραπεί σε CNF (βλ. βιβλίο J&M).
 - Η νέα γραμματική, όμως, μπορεί να μην οδηγεί πλέον στα συντακτικά δέντρα που θα θέλαμε.
- Ο αλγόριθμος CKY (ακολουθεί) είναι για ΓΧΣ σε CNF.
 - Πρόκειται για περίπτωση δυναμικού προγραμματισμού.

Αλγόριθμος CKY

① 0 θέλω ② 1 μία ③ 2 πρωινή ④ 3 πτήση ⑤ 4

	0	1	2	3	4
0		V (0,1) ↑	↑	↑	↑
1			Det (1,2) ↑	↑	↑
2				Adj (2,3) ↑	↑
3					Nominal N (3,4) ↑

Αλγόριθμος CKY

① 0 θέλω ② 1 μία ③ 2 πρωινή ④ 3 πτήση ⑤ 4

	0	1	2	3	4
0		V (0,1)	X (0,2)	Δεν υπάρχει κανόνας που να συνδυάζει V με Det.	
1			Det (1,2)		
2				Adj (2,3)	
3					Nominal N (3,4)

Αλγόριθμος CKY

① θέλω ② μία ③ πρωινή ④ πτήση ⑤

	0	1	2	3	4
0		V (0,1)			
1			Det (1,2)	X (1,3)	
2				Adj (2,3)	
3					Nominal N (3,4)

Δεν υπάρχει κανόνας που να συνδυάζει Det με Adj.

Αλγόριθμος CKY

① 0 θέλω ② 1 μία ③ 2 πρωινή ④ 3 πτήση ⑤ 4

	0	1	2	3	4
0		V (0,1)		X (0,3)	
1			Det (1,2)	(1,3)	
2				Adj (2,3)	
3					Nominal N (3,4)

Diagram illustrating the CKY algorithm table for the sentence "θέλω μία πρωινή πτήση". The table shows the decomposition of the sentence into phrases. The cell (0,3) is marked as empty (X) and highlighted in light blue. A red box contains the text "Το (1,3) είναι κενό." (The (1,3) is empty). Arrows indicate the flow of the algorithm: from (0,3) to (0,1) and from (0,3) to (1,3).

Αλγόριθμος CKY

(0) θέλω (1) μία (2) πρωινή (3) πτήση (4)

	0	1	2	3	4
0		V (0,1)	(0,2)	X (0,3)	
1			Det (1,2)	(1,3)	
2				Adj (2,3)	
3					Nominal N (3,4)

Το (0,2) είναι κενό.

Αλγόριθμος CKY

(0) θέλω (1) μία (2) πρωινή (3) πτήση (4)

	0	1	2	3	4
0		V (0,1)	(0,2)		
1			Det (1,2)	(1,3)	
2				Adj (2,3)	Nominal (2,4)
3					Nominal N (3,4)

A blue arrow points from the cell (2,4) to the cell (2,3). The word "Nominal" is written in green text next to the cell (2,4).

Αλγόριθμος CKY

① θέλω ② μία ③ πρωινή ④ πτήση

	0	1	2	3	4
0		V (0,1)	(0,2)		
1			Det (1,2)	(1,3)	NP (1,4)
2				Adj (2,3)	Nominal (2,4)
3					Nominal N (3,4)

Αλγόριθμος CKY

① 0 θέλω ② 1 μία ③ 2 πρωινή ④ 3 πτήση ⑤ 4

	0	1	2	3	4
0		V (0,1)	(0,2)	(0,3)	S (0,4) X X
1			Det (1,2)	(1,3)	NP (1,4)
2				Adj (2,3)	Nominal (2,4)
3					Nominal N (3,4)

Αλγόριθμος CKY

- Μπορούμε να **αποθηκεύουμε** σε κάθε κελί και τον κάθε **κανόνα που χρησιμοποιήθηκε**, καθώς και δείκτες προς τα **κελιά στα οποία εφαρμόστηκε** το δεξί μέρος του κανόνα.
 - Έτσι μπορούμε να **εξαγάγουμε** τελικά από τον πίνακα και το **συντακτικό δέντρο** (ή τα συντακτικά δέντρα αν είναι πολλά).
 - Αλλά η **εξαγωγή** του συντακτικού δέντρου κάνει **εκθετική** την **πολυπλοκότητα** του αλγορίθμου.
 - **Χωρίς εξαγωγή** $O(n^3)$ χρόνος, όπου n το πλήθος των λέξεων.
- Η **εξαγωγή** μπορεί να **επεκταθεί**, ώστε να επιστρέφει δέντρο της **αρχικής γραμματικής**.
 - Της γραμματικής που είχαμε πριν τη μετατροπή σε CNF.
 - Εναλλακτικά μπορούμε να χρησιμοποιήσουμε πιο περίπλοκους αλγορίθμους (π.χ. Earley) που δέχονται οποιαδήποτε ΓΧΣ.

Συμφωνία γένους με ΓΧΣ

- $S \rightarrow NP VP \mid VP$
- $NP \rightarrow Pron \mid \text{DetFem PNFem} \mid \text{DetFem NominalFem} \mid \text{DetMasc PNMasc} \mid \text{DetMasc NominalMasc}$
- $\text{NominalFem} \rightarrow \text{NFem} \mid \text{AdjFem NominalFem} \mid \text{NominalFem PP}$
- $\text{NominalMasc} \rightarrow \text{NMasc} \mid \text{AdjMasc NominalMasc} \mid \text{NominalMasc PP}$
- $VP \rightarrow V \mid V NP$
- $Pron \rightarrow \text{εγώ}$
- $\text{DetFem} \rightarrow \eta \mid \mu\iota\alpha \mid \tau\eta\nu$
- $\text{PNFem} \rightarrow \text{Θεσσαλονίκη} \mid \text{Αθήνα}$
- $\text{NFem} \rightarrow \text{πτήση}$
- $\text{AdjFem} \rightarrow \text{πρωινή} \mid \text{απογευματινή}$
- $V \rightarrow \text{θέλω} \mid \text{θέλει} \mid \text{προτιμώ} \mid \text{προτιμά}$
- $\text{Prep} \rightarrow \text{προς} \mid \text{από}$
- $PP \rightarrow \text{Prep NP}$
- $\text{DetMasc} \rightarrow \text{o} \mid \text{έναν} \mid \text{τον}$
- $\text{NMasc} \rightarrow \text{πελάτης} \mid \text{πελάτη}$

Διπλασιασμός των
σχετικών κανόνων. Και για
να πετύχουμε και συμφωνία
αριθμού και πτώσης;

Συμφωνία γένους με επαυξημένη ΓΧΣ

- $S \rightarrow NP VP \mid VP$
- $NP \rightarrow Pron \mid Det(\mathbf{G}) PN(\mathbf{G}) \mid Det(\mathbf{G}) Nominal(\mathbf{G})$
- $Nominal(\mathbf{G}) \rightarrow N(\mathbf{G}) \mid Adj(\mathbf{G}) Nominal(\mathbf{G}) \mid Nominal(\mathbf{G}) PP$
- $VP \rightarrow V \mid V NP$
- $PP \rightarrow Prep NP$
- $Pron \rightarrow \text{εγώ}$
- $Det(\mathbf{masc}) \rightarrow \text{ο} \mid \text{έναν} \mid \text{τον}$
- $Det(\mathbf{fem}) \rightarrow \text{η} \mid \text{μια} \mid \text{την}$
- $PN(\mathbf{fem}) \rightarrow \text{Θεσσαλονίκη} \mid \text{Αθήνα}$
- $N(\mathbf{fem}) \rightarrow \text{πτήση}$
- $N(\mathbf{masc}) \rightarrow \text{πελάτης} \mid \text{πελάτη}$
- $Adj(\mathbf{fem}) \rightarrow \text{πρωινή} \mid \text{απογευματινή}$
- $V \rightarrow \text{θέλω} \mid \text{θέλει} \mid \text{προτιμώ} \mid \text{προτιμά}$
- $Prep \rightarrow \text{προς} \mid \text{από}$

Αντίστοιχες ιδιότητες για
πτώση, αριθμό κλπ.

π.χ. $Det(\mathbf{fem}, \mathbf{nom}, \mathbf{sing}) \rightarrow \text{η}$

Δεν είναι πια ΓΧΣ, αλλά
μπορεί να μετατραπεί σε
ισοδύναμη ΓΧΣ με
περισσότερους κανόνες,
αρκεί οι δυνατές τιμές των
ιδιοτήτων να είναι
πεπερασμένες.

Οι επαυξημένες ΓΧΣ ως προτάσεις Horn

- Κάθε **ΓΧΣ** μπορεί να μετατραπεί σε **προτάσεις Horn**:
 - Π.χ. ο κανόνας $S \rightarrow NP VP$ γίνεται:
 $((NP(s_1) \wedge VP(s_2)) \Rightarrow S(s_1 + s_2))$, με «+» συνένωση συμβολοσειρών.
 - Αν η συμβολοσειρά s_1 είναι ένα NP και η συμβολοσειρά s_2 είναι ένα VP, τότε η συμβολοσειρά s_1+s_2 είναι ένα S.
 - Μπορούμε να εκτελέσουμε συντακτική ανάλυση με αλγορίθμους εξαγωγής συμπερασμάτων προτάσεων Horn.
- Για **επαυξημένη ΓΧΣ**, θα έχουμε **επιπλέον ορίσματα**:
 - Π.χ. ο $Nominal(G) \rightarrow Adj(G) Nominal(G)$ γίνεται:
 $((Adj(s_1, g) \wedge Nominal(s_2, g)) \Rightarrow Nominal(s_1 + s_2, g))$,
όπου g μεταβλητή
 - και ο $N(masc) \rightarrow$ πελάτης γίνεται:
 $N("πελάτης", Masc)$.

Εκτός
εξεταστέας
ύλης.

Συντακτική ανάλυση με Prolog

- Η **Prolog** υποστηρίζει άμεσα **γραμματικές DCG**.
 - **Definite Clause Grammars. ΓΧΣ** επαυξημένες με ιδιότητες, όπως στο παράδειγμα συμφωνίας γένους.
 - Τις μετατρέπει σε **προτάσεις Horn ΠΚΛ** (για την ακρίβεια οριστικές προτάσεις – definite clauses) και αντιμετωπίζει τη συντακτική ανάλυση ως **πρόβλημα εξαγωγής συμπεράσματος** (βλ. μάθημα «Τεχνητή Νοημοσύνη»).
 - Μπορούμε να δώσουμε **απευθείας** τις γραμματικές στο μεταγλωττιστή/διερμηνέα της Prolog.
- Στην πράξη χρησιμοποιούνται συνήθως **πιο σύνθετοι** αλγόριθμοι συντακτικής ανάλυσης.
 - Π.χ. **CKY, Earley**, ενδεχομένως **τροποποιημένοι**, ώστε να δέχονται **επαυξημένες γραμματικές** χωρίς συμφραζόμενα.
 - Μπορούν να **υλοποιηθούν** και αυτοί σε Prolog ή σε άλλες γλώσσες προγραμματισμού.

Οι γραμματικές DCG της Prolog

- ΓΧΣ επαυξημένες με ιδιότητες.
 - Π.χ. `nominal(G) → adj(G), nominal(G)`.
 - Π.χ. `det(masc) → [έναν]`.
 - Τα **τερματικά** σύμβολα γράφονται μέσα σε **αγκύλες**.
 - Αν το πρώτο γράμμα είναι **κεφαλαίο**, πρόκειται για **μεταβλητή**.
- **Περιορισμός** από τον ενσωματωμένο αλγόριθμο DFS:
 - Πρέπει να αποφεύγουμε κανόνες με **αριστερή αναδρομή**.
 - Π.χ. `nominal → nominal, pp`.
 - Γενικότερα πρέπει να αποφεύγουμε γραμματικές που επιτρέπουν παραγωγές της μορφής $A \rightarrow \dots \rightarrow A \dots$

Παράδειγμα γραμματικής DCG

s --> np, vp.

s --> vp.

np --> pron.

np --> det(G), pn(G).

np --> det(G), nominal(G).

nominal(G) --> n(G).

nominal(G) --> adj(G), nominal(G).

% Πρόβλημα αριστερής αναδρομής:

% nominal(G) --> nominal(G), pp.

nominal(G) --> n(G), manypp.

manypp --> pp.

manypp --> pp, manypp.

vp --> v.

vp --> v, np.

pp --> prep, np.

pron --> [εγώ].

det(masc) --> [ο].

det(masc) --> [έναν].

...

Περισσότερα
παραδείγματα
γραμματικών DCG στα
έγγραφα του μαθήματος
στο e-class.

Χρήση γραμματικών DCG

- Θα χρειαστείτε μεταγλωττιστή/διερμηνέα **Prolog**.
 - Π.χ. **SWI-Prolog** (<https://www.swi-prolog.org/>).
- **Φόρτωμα** του αρχείου της γραμματικής:
 - **consult(...)** στη γραμμή εντολών της Prolog.
 - Σε Windows: διπλό κλικ στο αρχείο .pl της γραμματικής.
 - Υπάρχουν παραδείγματα γραμματικών στις ιστοσελίδες.
- Συντακτική **ανάλυση**:
 - **phrase(s, [θέλω, μια, πτήση, από, την, αθήνα])**.
 - **phrase(nominal(masc), [πελάτης, από, την, αθήνα])**.
 - Η απόκριση **yes/no** της Prolog σημαίνει ότι υπάρχει ή όχι συντακτικό δέντρο με αυτή τη ρίζα.

Χρήση γραμματικών DCG – συνέχεια

- **Ερωτήσεις** προς το συντακτικό αναλυτή:
 - **phrase(nominal(G), [πελάτης, από, την, αθήνα]).**
 - Απάντηση: **G = masc.**
 - Με «;» ζητάμε άλλη απάντηση (εδώ δεν υπάρχει).
- **Επιστροφή συντακτικού δένδρου:**
 - Μπορούμε να επεκτείνουμε τη γραμματική (βλ. παρακάτω προαιρετικές διαφάνειες), ώστε να επιστρέφεται και συντακτικό δέντρο:
 - **phrase(nominal(G, T), [πελάτης, από, την, αθήνα]).**
 - Απάντηση: **G = masc** και:
 - **T = nominal(n(πελάτης),
 manypp(pp(prepare(από),
 np(det(την),
 pn(αθήνα))))))**

Κόμβοι με παραστάσεις υποδένδρων

Εκτός
εξεταστέας
ύλης.

nominal (fem, **nominal(adj(πρωινή), nominal(n(πτήση)))**)

Σε κάθε κόμβο του
συντακτικού δέντρου
έχουμε ένα επιπλέον
όρισμα που παριστάνει
το υποδέντρο που έχει
ρίζα αυτόν τον κόμβο.

nominal(fem, **nominal(n(πτήση))**)

adj(fem, **adj(πρωινή)**)

n(fem, **n(πτήση)**)

πρωινή

πτήση

Νέα μορφή των κανόνων

$\text{adj}(\text{fem}, \text{adj}(\text{πρωινή})) \rightarrow [\text{πρωινή}]$.

$\text{n}(\text{fem}, \text{n}(\text{πτήση})) \rightarrow [\text{πτήση}]$.

$\text{n}(\text{masc}, \text{n}(\text{πελάτης})) \rightarrow [\text{πελάτης}]$.

$\text{nominal}(\text{G}, \text{nominal}(\text{T})) \rightarrow \text{n}(\text{G}, \text{T})$.

$\text{nominal}(\text{G}, \text{nominal}(\text{T1}, \text{T2})) \rightarrow \text{adj}(\text{G}, \text{T1}), \text{nominal}(\text{G}, \text{T2})$.

Αν βρεις ένα επίθετο (adj) γένους G με συντακτικό δέντρο T1, ακολουθούμενο από ένα nominal γένους G με συντακτικό δέντρο T2, τότε έχεις βρει ένα (μεγαλύτερο) nominal γένους G με συντακτικό δέντρο nominal(T1, T2).

Δείτε το αρχείο *tree_structure.pl* στις ιστοσελίδες του μαθήματος.

DCG για απλή γλώσσα αριθμητικής

digit --> [zero].

digit --> [one].

...

digit --> [nine].

expression --> digit.

expression --> [open], expression, [plus], expression, [close].

expression --> [open], expression, [minus], expression, [close].

expression --> [open], expression, [star], expression, [close].

expression --> [open], expression, [slash], expression, [close].

open open two plus four close slash
open four minus one close close

$((2 + 4) / (4 - 1))$

➤ phrase(expression, [open, open, two, plus, four, close, slash,
open, four, minus, one, close, close]).

Yes.

Σημασιολογία αριθμητικής γλώσσας

digit(0) --> [zero].

digit(1) --> [one].

...

digit(9) --> [nine].

Μέσα σε άγκιστρα γράφουμε επιπλέον περιορισμούς που πρέπει να ικανοποιούνται για να χρησιμοποιηθεί ο κανόνας. Εδώ με το is εκχωρούμε στο X το αποτέλεσμα του αριθμητικού υπολογισμού $X1 + X2$. (Το = στην Prolog σημαίνει ενοποίηση.)

expression(X) --> digit(X).

expression(X) --> [open], expression(X1), [plus], expression(X2),
[close], {X is X1 + X2}.

expression(X) --> [open], expression(X1), [minus],
expression(X2), [close], {X is X1 - X2}.

...

➤ phrase(expression(X), [open, open, two, plus, four, close, slash,
open, four, minus, one, close, close]).

X = 2.

DCG για κλείσιμο εισιτηρίων

- Απλοϊκή γραμματική που επιστρέφει τη **σημασιολογική παράσταση** αιτημάτων **κλεισίματος εισιτηρίων**.

- **phrase(s(Semantics), [επιθυμώ, μια, πτήση, από, την, αθήνα, προς, τη, Θεσσαλονίκη]).**

- Απάντηση:

Semantics = request(speaker, buy,
[ticketType(flight),
location(departure, ath),
location(arrival, thess)])

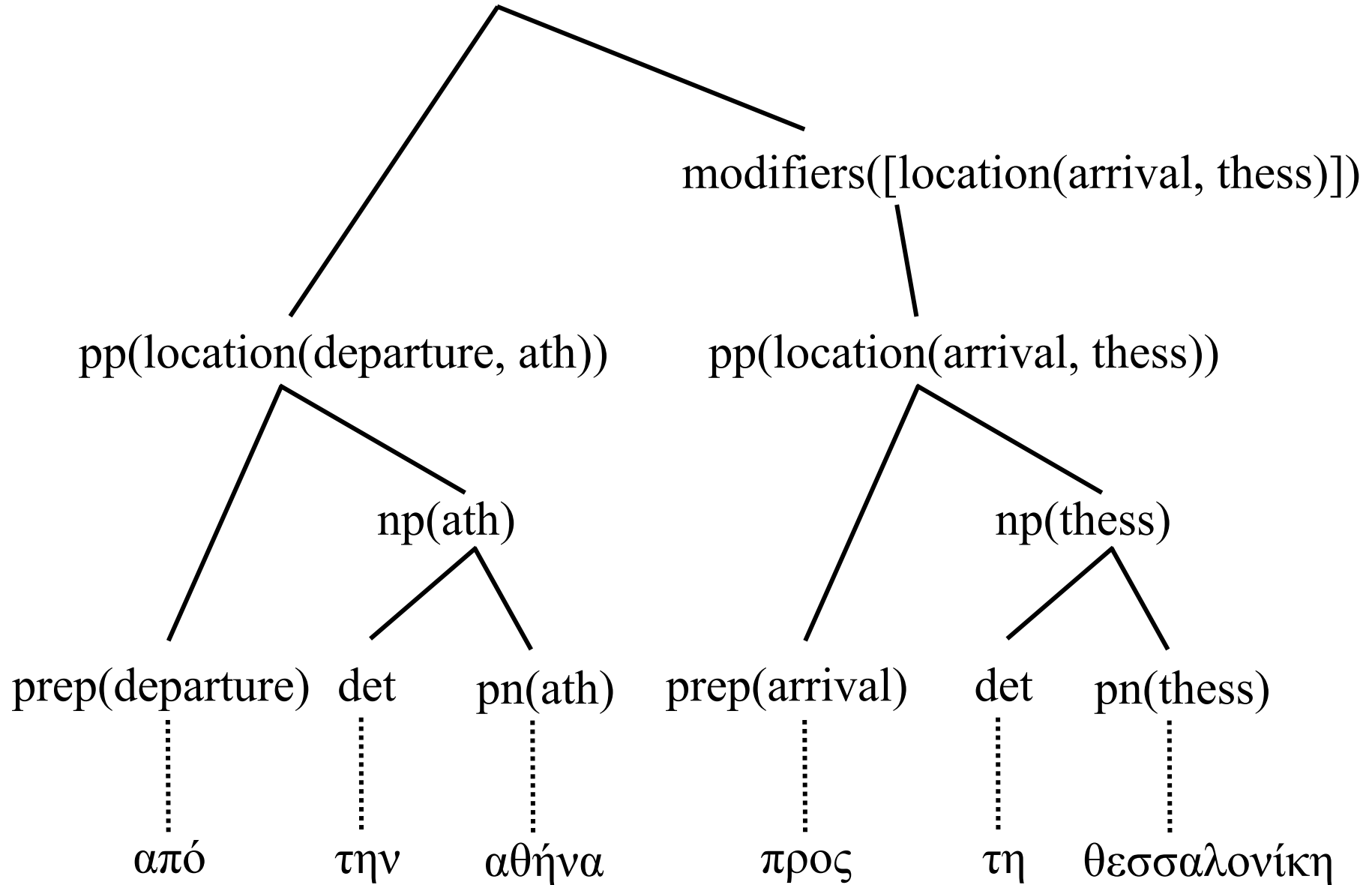
- **phrase(s(Semantics), [εγώ, θέλω, ένα, τρένο, προς, τη, Θεσσαλονίκη, αύριο]).**

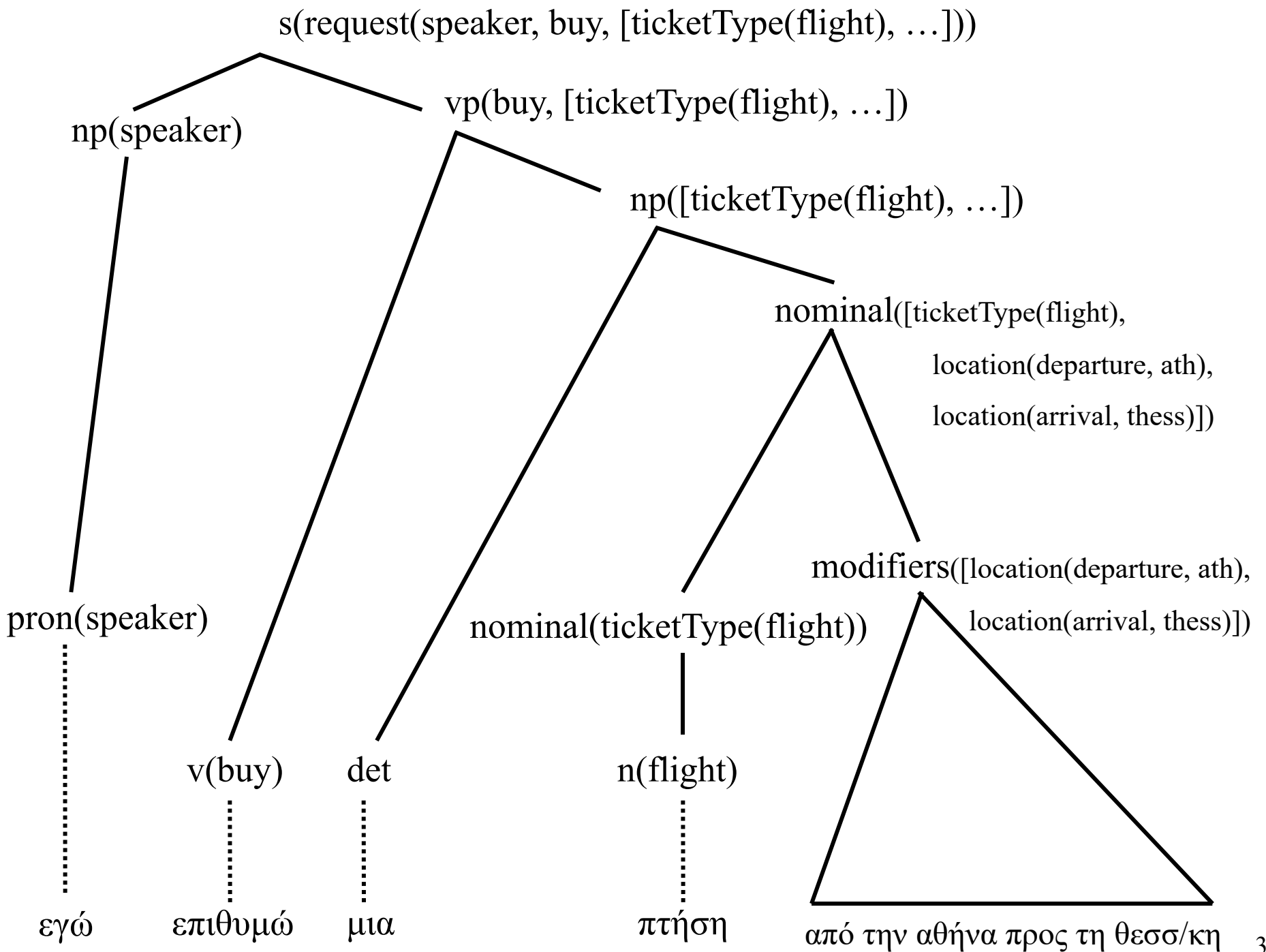
- Απάντηση:

Semantics = request(speaker, buy,
[date(tomorrow), ticketType(train),
location(arrival, thess)])

Τι θέλουμε να κάνει η γραμματική

modifiers([location(departure, ath), location(arrival, thess)])





Κύρια ονόματα, ουσιαστικά, άρθρα, αντωνυμίες

- Κάθε κύριο όνομα (π.χ. «αθήνα») αντιστοιχίζεται στον αντίστοιχο κωδικό της πόλης (π.χ. «ath»).
- $pn(ath) \rightarrow [αθήνα].$ $pn(thess) \rightarrow [θεσσαλονίκη].$
 $pn(herak) \rightarrow [ηράκλειο].$
- Τα μόνα ουσιαστικά που είναι δυνατόν να χρησιμοποιηθούν δείχνουν αν ο πελάτης ενδιαφέρεται για αγορά αεροπορικού ή σιδηροδρομικού εισιτηρίου.
- $n(flight) \rightarrow [πτήση].$
 $n(train) \rightarrow [τρένο].$
- Στο απλοϊκό αυτό παράδειγμα γραμματικής, τα άρθρα δεν συνεισφέρουν στη σημασιολογική παράσταση.
- $det \rightarrow [τη].$ $det \rightarrow [την].$ $det \rightarrow [το].$...
- Και έχουμε μόνο μία αντωνυμία.
- $pron(speaker) \rightarrow [εγώ].$

Ρήματα, επιρρήματα, προθέσεις

- Στο απλοϊκό αυτό παράδειγμα γραμματικής, τα **ρήματα** έχουν όλα την ίδια σημασία.
v(buy)--> [θέλω].
v(buy) --> [επιθυμώ].
- Σε αυτή τη γραμματική, υπάρχουν μόνο **χρονικά επιρρήματα**, που δείχνουν την **ημέρα αναχώρησης**.
adv(today) --> [σήμερα].
adv(tomorrow) --> [αύριο].
- Στη γραμματική αυτή, οι **προθέσεις** δείχνουν αν η τοποθεσία που ακολουθεί αποτελεί το **σημείο αναχώρησης** ή τον **προορισμό**.
prep(departure) --> [από].
prep(arrival) --> [προς].

Ονοματική φράση

- Εδώ θεωρούμε ότι μια ονοματική φράση είναι ένα **κύριο όνομα** (pn) με άρθρο ή ένα **nominal** με το άρθρο του ή μια αντωνυμία.

άρθρο + κύριο όνομα: «τη θεσσαλονίκη»

άρθρο + nominal: «μια πτήση προς τη θεσσαλονίκη»

αντωνυμία: «εγώ»

- Η **σημασιολογική παράσταση** της ονοματικής φράσης είναι η σημασιολογική παράσταση **του κυρίου ονόματος ή του nominal ή της αντωνυμίας** αντίστοιχα.

np(**PNSem**) --> det, **pn(PNSem)**.

np(**NomSem**) --> det, **nominal(NomSem)**.

np(**PronSem**) --> pron(**PronSem**).

Εμπρόθετοι προσδιορισμοί

- Σε αυτή τη γραμματική, ένας **εμπρόθετος προσδιορισμός** (prepositional phrase, pp) είναι μια **πρόθεση** ακολουθούμενη από μια **ονοματική φράση**.
 - Π.χ. «**από** **την αθήνα**», «**προς** **τη θεσσαλονίκη**».
 - Η σημασιολογική παράσταση της **ονοματικής φράσης** προσδιορίζει μία πόλη (π.χ. **ath**).
 - Η σημασιολογική παράσταση της **πρόθεσης** δείχνει αν η πόλη είναι προορισμός ή σημείο αναχώρησης (**departure** ή **arrival**).
 - Η **σημασιολογική παράσταση** του εμπρόθετου προσδιορισμού είναι ένα κατηγορημα **location**.
 - Δείχνει αν προσδιορίζεται ο **προορισμός** ή το **σημείο αναχώρησης** και **ποιος είναι αυτός**.
- pp(location(PrepSem, NPSem)) -->**
prep(PrepSem), np(NPSem).

Πολλοί εμπρόθετοι προσδιορισμοί

- Το **modifiers** είναι ένας ή περισσότεροι εμπρόθετοι προσδιορισμοί (**pp**).
- Αν έχουμε **μόνο ένα** pp, η σημασιολογική παράσταση του modifiers είναι μια **λίστα** με την παράσταση του pp.
- Διαφορετικά, η σημασιολογική παράσταση του modifiers είναι μια **λίστα** που περιλαμβάνει τις σημασιολογικές παραστάσεις των pp.

– Π.χ. **[location(departure, ath), location(arrival, thess)]**

- Οι σχετικοί κανόνες της γραμματικής:

modifiers([PPSem]) --> pp(PPSem).

modifiers([PPSem | ModSem]) -->
pp(PPSem), modifiers(ModSem).

Το **[A | L]** φτιάχνει ένα αντίγραφο της λίστας L που έχει το πρόσθετο στοιχείο A στην αρχή της.

Οι κανόνες για το nominal

- Εδώ ένα nominal είναι ένα **ουσιαστικό**, πιθανώς συνοδευόμενο από **εμπρόθετους προσδιορισμούς**.
 - Π.χ. «**τρένο** από την αθήνα προς τη θεσσαλονίκη»
- Η σημασιολογική παράσταση του **ουσιαστικού** (**flight** ή **train**) γίνεται το όρισμα ενός κατηγορήματος **ticketType**.
 - Π.χ. **ticketType(train)**
- Αν υπάρχουν **εμπρόθετοι προσδιορισμοί (modifiers)**, το κατηγορήμα **ticketType** προστίθεται στην αρχή της **λίστας** με τις **παραστάσεις των εμπρόθετων προσδιορισμών**.
 - Π.χ. [**ticketType(train)**,
and(location(departure, ath), location(arrival, thess))]
- Οι κανόνες της γραμματικής:
nominal(ticketType(NSem)) --> n(NSem).
nominal([NomSem | ModSem]) -->
nominal(NomSem), modifiers(ModSem).

Ρηματικές φράσεις

- Στη γραμματική αυτή, η ρηματική φράση περιλαμβάνει ένα **μεταβατικό ρήμα**, μια **ονοματική φράση** (το αντικείμενο του ρήματος) και ενδεχομένως ένα **χρονικό επίρρημα** (με παράσταση «**today**» ή «**tomorrow**»).
- Π.χ. «**θέλω** μια πτήση»
- Π.χ. «**θέλω** μια πτήση προς τη θεσσαλονίκη»
- Π.χ. «**θέλω** μια πτήση προς τη θεσσαλονίκη **αύριο**»
- Οι αντίστοιχοι κανόνες της γραμματικής:
$$\text{vp}(\text{VSem}, \text{NPSem}) \rightarrow \text{v}(\text{VSem}), \text{np}(\text{NPSem}).$$
$$\text{vp}(\text{VSem}, [\text{date}(\text{AdvSem}) \mid \text{NPSem}]) \rightarrow \text{v}(\text{VSem}), \text{np}(\text{NPSem}), \text{adv}(\text{AdvSem}).$$

Προτάσεις

- Στη γραμματική αυτή, η πρόταση (s) αποτελείται είτε **μόνο από μία ρηματική φράση (vr)** είτε από **μία ονοματική φράση (nr)** και **μία ρηματική φράση (vr)**.
 - Π.χ. «θέλω μια πτήση προς τη Θεσσαλονίκη».
 - Π.χ. «εγώ θέλω μια πτήση προς τη Θεσσαλονίκη».
- Οι αντίστοιχοι κανόνες της γραμματικής:
 $s(\text{request}(\text{speaker}, \text{VSem}, \text{ObjSem})) \rightarrow \text{vr}(\text{VSem}, \text{ObjSem}).$
 $s(\text{request}(\text{NPsem}, \text{VSem}, \text{ObjSem})) \rightarrow$
 $\text{nr}(\text{NPsem}), \text{vr}(\text{VSem}, \text{ObjSem}).$

Διάβασμα

- Το μεγαλύτερο μέρος της ύλης αυτής της ενότητας καλύπτεται από τα κεφάλαια 12 και 13 των D. Jurafsky and J.H. Martin, 2^η έκδοση, Prentice Hall, 2008.
 - Υπάρχει στη βιβλιοθήκη του ΟΠΑ.
 - Εισαγωγή κεφ. 12, ενότητες 12.1–12.3, 12.5, 13.2, 13.4.1.
 - Για τις εξετάσεις, χρειάζεται να γνωρίζετε μόνο όσα αναφέρουν οι διαφάνειες.
 - Όσοι ενδιαφέρονται μπορούν να διαβάσουν προαιρετικά και το κεφάλαιο 18.
- Μπορείτε να συμβουλευτείτε και την 3^η έκδοση του βιβλίου των Jurafsky & Martin (υπό προετοιμασία), που διατίθεται δωρεάν.
 - <http://web.stanford.edu/~jurafsky/slp3/>.

