

Ερώτηση: 1 (Βαθμός: 0.50 / 1.00)
(Πολυαπάνης Επιδότησης (Μοναδική Απάντηση))

Υπολογιστική Γλωσσολογία ονομάζεται η επιστημονική περιοχή μεταξύ:

Απάντηση	
<input checked="" type="checkbox"/> ✖	Γλωσσολογίας και Ιατρικής (λάθος)
<input type="checkbox"/>	Ιατρικής και Πληροφορικής (λάθος)
<input type="checkbox"/>	Πληροφορικής και Γλωσσολογίας (σωστό)

Η κωδικοποίηση Unicode...	
Απάντηση	<div><input checked="" type="checkbox"/> ✔ Μπορεί να αναπαράσχει πάνω από 1.000.000 διαφορετικούς χαρακτήρες. (σωστό)</div> <div><input type="checkbox"/> Περιλαμβάνει 17 επίπεδα με πάνω από 60.000 code points το καθένα. (σωστό)</div> <div><input type="checkbox"/> Μπορεί να χρησιμοποιηθεί για την σύνοψη/χαρακτήρων από περισσότερες από δύο γλώσσες στο ίδιο κείμενο. (σωστό)</div> <div><input type="checkbox"/> Αφορά μόνο αριθμητικά ψηφία. (λάθος)</div> <div><input type="checkbox"/> Είναι παλαιότερη από την κωδικοποίηση ASCII. (λάθος)</div>

(Σχεδόν) συνώνυμοι όροι με την Υπολογιστική Γλωσσολογία είναι οι εξής:

Απάντηση	
<input checked="" type="checkbox"/> ✖	Γλωσσικός Υπολογισμός (λάθος)
<input type="checkbox"/>	Επεξεργασία Φυσικής Γλώσσας (σωστό)
<input type="checkbox"/>	Φυσική Επεξεργασία Γλώσσας (λάθος)
<input type="checkbox"/>	Γλωσσική Τεχνολογία (σωστό)
<input type="checkbox"/>	Τεχνητή Γλωσσολογία (λάθος)

Ο δεκαεξαδικός αριθμός D8 στον κώδικα ISO 8859-7 αντιστοιχεί στον χαρακτήρα:

Απάντηση	
<input checked="" type="checkbox"/> ✖	X (λάθος)
<input type="checkbox"/>	Θ (λάθος)
<input type="checkbox"/>	Ω (λάθος)
<input type="checkbox"/>	= (λάθος)
<input type="checkbox"/>	Κανένα από τα παραπάνω. (σωστό)

ACL είναι τα αρχικά του οργανισμού, το πλήρες όνομα του οποίου είναι:

Απάντηση	
<input checked="" type="checkbox"/> ✖	American Computational Linguistics (λάθος)
<input type="checkbox"/>	Association for Computational Linguistics (σωστό)
<input type="checkbox"/>	Association for Computer Languages (λάθος)
<input type="checkbox"/>	Administration for Computer Languages (λάθος)

Η κωδικοποίηση των Αγγλικών χαρακτήρων είναι συμβατός τόσο στον κώδικα ASCII και στους κώδικες ISO 8859 και στους κώδικες ANSI (Windows).

Απάντηση	
<input checked="" type="checkbox"/> ✔	Σωστό (σωστό)
<input type="checkbox"/>	Λάθος (λάθος)

Ο χαρακτήρας F στον κώδικα ASCII έχει δεκαεξαδική αναπαράσταση:

Απάντηση

☒

46

(σωστό)

☐

64

(λάθος)

☐

F

(λάθος)

☐

7F

(λάθος)

☐

Κανένα από τα παραπάνω.

(λάθος)

Στην αρχή του κώδικα ISO 8859, υπάρχει ο κώδικας ASCII.

Απάντηση

☒

Σωστό (σωστό)

☐

Λάθος (λάθος)

Στην αρχή της κωδικοσελίδας CP 1253, υπάρχει ο κώδικας ASCII.

Απάντηση

☒

Σωστό (σωστό)

☐

Λάθος (λάθος)

Η απόσταση Levenshtein μεταξύ των λέξεων ΚΑΛΥΓΑ και ΑΚΥΓΛΑ είναι:

Απάντηση

☒

1

(λάθος)

☐

2

(λάθος)

☐

3

(σωστό)

☐

4

(λάθος)

☐

5

(λάθος)

Ο δεκαεξαδικός αριθμός 65 στον κώδικα ASCII αντιστοιχεί στον χαρακτήρα:

Απάντηση

☐

F

(λάθος)

☒

\$

(λάθος)

☐

x

(λάθος)

☐

e

(σωστό)

☐

Κανένα από τα παραπάνω.

(λάθος)

Το πλήθος των βλάντων ποσειτών στον πίνακα Levenshtein για το υπολογισμό της απόστασης μεταξύ των λέξεων ΗΛΙΚΙΑ και ΛΟΔΙΑ είναι:

Απάντηση

☒

1

(λάθος)

☐

2

(σωστό)

☐

3

(λάθος)

☐

4

(λάθος)

☐

Μεγαλύτερο από 4.

Το πλῆθος των βέλτιστων μοναδικών στον τίτλο Levenshtein για τον υπολογισμό της απόστασης μεταξύ των λέξεων γινάα και ιηκίαα είναι:

Απάντηση

✖

✓

1

(σωστό)

☐

2.

(λάθος)

☐

3.

(λάθος)

☐

4.

(λάθος)

☐

Μεγλύτερο από 4.

(λάθος)

Έστω η κανονική έκφραση: $a|(bc)(cb)a|$

Ποιές από τις παρακάτω συμβολοσειρές περιλαμβάνονται στη γλώσσα που ορίζει;

Απάντηση

☒

✖

abc

(σωστό)

☐

aba

(λάθος)

☐

acb

(σωστό)

☐

aca

(λάθος)

☐

cba

(λάθος)

☐

aa

(σωστό)

Έστω δυο λέξεις A και B που απέρχουν 3 μονάδες (πλήθος διαφορετικών) σύμφωνα με την απόσταση Levenshtein, και μια τρίτη λέξη Γ που απέρχεται από την A. Πόση μπορεί να απέρχεται Γ από την A;

Απάντηση

☒

✖

4

(λάθος)

☐

Μεταξύ 3 και 5

(λάθος)

☐

Μεταξύ 2 και 8

(σωστό)

☐

Κανένα από τα παραπάνω

(λάθος)

Έστω η κανονική έκφραση: $//ab?Hca!?$

Ποιές από τις παρακάτω συμβολοσειρές περιλαμβάνονται στη γλώσσα που ορίζει;

Απάντηση

☒

✖

bba

(λάθος)

☐

aba

(σωστό)

☐

cacaca

(λάθος)

☐

abca

(σωστό)

☐

ababcca

(σωστό)

☐

aabc

(λάθος)

Έστω δυο λέξεις A και B που απέρχουν 4 μονάδες (πλήθος διαφορετικών) σύμφωνα με την απόσταση Levenshtein, και μια τρίτη λέξη Γ που απέρχεται από την A. Πόση μπορεί να απέρχεται Γ από την A;

Απάντηση

☒

✖

Μεταξύ 3 και 5

(σωστό)

☐

Μεταξύ 1 και 4

(λάθος)

☐

6

(λάθος)

☐

Κανένα από τα παραπάνω

(λάθος)

Έστω η κανονική έκφραση: $a|(bc)?b|bc|?$

Ποιές από τις παρακάτω συμβολοσειρές περιλαμβάνονται στη γλώσσα που ορίζει;

Απάντηση

☒

✖

abcb

(λάθος)

☐

abcc

(σωστό)

☐

abccb

(λάθος)

☐

abcbbb

(σωστό)

☐

abcbc

(σωστό)

☐

abb

(σωστό)

Έστω η κανονική έκφραση: /abc)*a(b)c)?/

Ποιές από τις παρακάτω συμβολοσειρές περιλαμβάνονται στη γλώσσα που ορίζει:

Απάντηση

☒ ✖

abcac

(λάθος)

☐

abcabc

(σωστό)

☐

aabcc

(λάθος)

☐

abcab

(σωστό)

☐

aababc

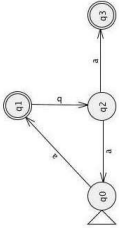
(λάθος)

☐

abcabcc

(σωστό)

Το παρακάτω αυτόματο είναι:



Απάντηση

☒ ✔

Κυβλιώ

(σωστό)

☐

Ακυβλιώ

(λάθος)

☐

Απογορευτικό

(λάθος)

☐

Μη-απογορευτικό

(σωστό)

Υπάρχουν δύο μεταβάσεις από την q2 με το ίδιο σύμβολο.

Επιλέξτε όσες από τις παρακάτω κανονικές εκφράσεις αναγνωρίζουν τη συμβολοσειρά: aabbb

Απάντηση

☒ ✔

/ab)+b/

(σωστό)

☐

/c*(ab)*b/

(σωστό)

☐

/ab*ab+*/

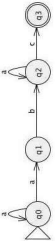
(σωστό)

☐

/a(ba)*b/

(λάθος)

Με ποιά από τις παρακάτω κανονικές εκφράσεις είναι ισοδύναμο το ακόλουθο αυτόματο πεπερασμένων καταστάσεων:



Απάντηση

☒ ✖

/a+b*/

(λάθος)

☐

/a*ba*b*/

(λάθος)

☐

/a*ba+c/

(λάθος)

☐

/a+b*b*/

(σωστό)

Το παρακάτω αυτόματο είναι:



Απάντηση

☒ ✔

Κυβλιώ

(σωστό)

Υπάρχει μια κυβλική μετάβαση από την κατάσταση q2 στην κατάσταση q2.

☐

Ακυβλιώ

(λάθος)

☐

Απογορευτικό

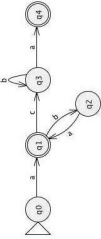
(σωστό)

☐

Μη-απογορευτικό

(λάθος)

Με ποιά από τις παρακάτω κανονικές εκφράσεις είναι ισοδύναμο το ακόλουθο αυτόματο πεπερασμένων καταστάσεων:



Απάντηση

☒ ✔

/abba*(c*a)*b/

(σωστό)

☐

/abba*b*a*/

(λάθος)

☐

/abba*(c)a*/

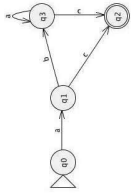
(λάθος)

☐

/abba*(c*a)*b/

(λάθος)

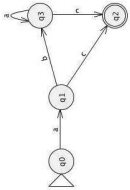
Ο πίνακας που αντιστοιχεί στο παρακάτω αυτόματο έχει διατυπώσει:



Απάντηση	
<input checked="" type="checkbox"/> X	3 γραμμές, 4 στήλες (0,000)
<input type="checkbox"/>	4 γραμμές, 3 στήλες (0,0000)
<input type="checkbox"/>	4 γραμμές, 4 στήλες (0,0000)
<input type="checkbox"/>	3 γραμμές, 3 στήλες (0,0000)

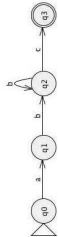
Οι γραμμές είναι όλες οι καταστάσεις. Οι στήλες είναι όλες τα σύμβολα που λαμβάνει.

Με ποιά από τις παρακάτω κανονικές εκφράσεις είναι ισοδύναμο το ακόλουθο αυτόματο πεπερασμένων καταστάσεων;



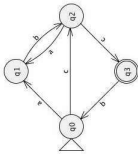
Απάντηση	
<input checked="" type="checkbox"/> X	$a^2(b^2+ac)^*$ (0,0000)
<input type="checkbox"/>	$a^2b^2bac^2$ (0,0000)
<input type="checkbox"/>	$a^2+caabac^2$ (0,0000)
<input type="checkbox"/>	$a^2(ba^2+c)^*$ (0,0000)

Με ποιά από τις παρακάτω κανονικές εκφράσεις είναι ισοδύναμο το ακόλουθο αυτόματο πεπερασμένων καταστάσεων;



Απάντηση	
<input checked="" type="checkbox"/> X	a^2b^2+c (0,0000)
<input type="checkbox"/>	a^2b^2c (0,0000)
<input type="checkbox"/>	a^2b^2+c (0,0000)
<input type="checkbox"/>	a^2b^2c+c (0,0000)

Ποιές από τις παρακάτω συμβολοσειρές ανηγούνται, το ακόλουθο αυτόματο πεπερασμένων καταστάσεων;



Απάντηση	
<input checked="" type="checkbox"/> X	ca^2bab (0,0000) Δεν ανηγείται σε κάποια κατάσταση
<input type="checkbox"/>	$ababac$ (0,0000)
<input type="checkbox"/>	ca^2bab (0,0000)
<input type="checkbox"/>	ca^2bc (0,0000)
<input type="checkbox"/>	ab^2c (0,0000)
<input type="checkbox"/>	ab^2babac (0,0000)

Ο ακόλουθος πίνακας φέρει ένα αυτόματο πεπερασμένων καταστάσεων. Ποιές από τις παρακάτω συμβολοσειρές παραλαμβάνονται στη γλώσσα που φέρει;

RSA			
>1	a	b	c
2	2	-	-
2	-	3,5	4
3	-	-	5
4:	-	5	-
5:	-	-	-

Απάντηση	
<input checked="" type="checkbox"/> X	ac (0,0000)
<input type="checkbox"/>	ab ² c (0,0000)
<input type="checkbox"/>	ab ² (0,0000)
<input type="checkbox"/>	abba (0,0000)
<input type="checkbox"/>	abcc (0,0000)
<input type="checkbox"/>	acaa (0,0000)

Επιλέξτε τον σωστό τύπο για τον κανόνα επαναγγραφής $Q \rightarrow a \vee B \vee a$

Απάντηση

☒ **x**

0

(λάθος)

Τόσο το κανόνας τύπου $Q \vee a$ τύπου B , τσαντά με το διάστημα να συμπεριλάβει ο μεγαλύτερος αριθμός τύπου που αντιστοιχεί στον κανόνα.

☐ 1

(σωστό)

Τόσο το κανόνας τύπου $Q \vee a$ τύπου B , τσαντά με το διάστημα να συμπεριλάβει ο μεγαλύτερος αριθμός τύπου που αντιστοιχεί στον κανόνα.

☐ 2

(σωστό)

Το a στο δεξιά της γραμμής του Q αποδοθούν "αναγράφονται".

☐ 3

(λάθος)

Σύμφωνα με την Ιεραρχία Chomsky, οι γραμμικές, Τύπου 3 επιτρέπονται στις γραμμικές, Τύπου 1.

Απάντηση

☒ **✓**

Σωστό (σωστό)

☐ Λάθος (λάθος)

Επιλέξτε τον σωστό τύπο για τον κανόνα επαναγγραφής $a \vee b \rightarrow a \vee B \vee a$

Απάντηση

☒ 0

(σωστό)

☐ 1

(λάθος)

☐ 2

(λάθος)

☐ 3

(λάθος)

Συντακτικά, οι φυσικές γλώσσες μπορούν να εκφραστούν ικανοποιητικά με γραμμικούς τύπου 2.

Απάντηση

☒ **✓**

Σωστό (σωστό)

Ναι, με ελάχιστη έκταση επάνω συντακτικά φαινόμενα που απαιτούν γραμμική τύπου 1.

☐ Λάθος (λάθος)

Επιλέξτε τον σωστό τύπο για τον κανόνα επαναγγραφής $a \vee Q \rightarrow a \vee B \vee a$

Απάντηση

☒ **x**

0

(λάθος)

Τόσο το κανόνας τύπου $a \vee Q$ τύπου B , τσαντά με το διάστημα να συμπεριλάβει ο μεγαλύτερος αριθμός τύπου που αντιστοιχεί στον κανόνα.

☐ 1

(σωστό)

Το a στο δεξιά της γραμμής του Q αποδοθούν "αναγράφονται".

☐ 2

(λάθος)

☐ 3

(λάθος)

Επιλέξτε τη σωστή μορφή CNF που αντιστοιχεί στον κανόνα: $W \rightarrow a \vee b \vee c$

Απάντηση

☒ **x**

$W \rightarrow a \vee B \vee C$

$A \rightarrow a$

$B \rightarrow b$

$C \rightarrow c$

(λάθος)

☐

$W \rightarrow D \vee E$

$D \rightarrow A \vee B$

$E \rightarrow C$

$A \rightarrow a$

$B \rightarrow b$

(λάθος)

☐

$W \rightarrow D \vee C$

$D \rightarrow A \vee B$

$A \rightarrow a$

$B \rightarrow b$

$C \rightarrow c$

(σωστό)

☐

Είναι ήδη σε μορφή CNF.

(λάθος)

Επιλέξτε τη σωστή μορφή CNF που αντιστοιχεί στον κανόνα: $C \rightarrow A \vee B$

Απάντηση	
<input checked="" type="checkbox"/> ✖	$C \leftrightarrow a$ $C \leftrightarrow b$ $A \leftrightarrow a$ $B \leftrightarrow b$ (λάθος)
<input type="checkbox"/>	$C \leftrightarrow a$ $C \leftrightarrow b$ (λάθος)
<input type="checkbox"/>	$C \leftrightarrow a \vee b$ (λάθος)
<input type="checkbox"/>	Είναι ήδη σε μορφή CNF. (σωστό)

Επιλέξτε τη σωστή μορφή CNF που αντιστοιχεί στον κανόνα: $C \rightarrow B$

Απάντηση	
<input type="checkbox"/>	$C \leftrightarrow b$ (λάθος)
<input checked="" type="checkbox"/> ✖	$C \leftrightarrow b$ $B \leftrightarrow b$ (λάθος)
<input type="checkbox"/>	Πρέπει να γνωρίζουμε και τους υπόλοιπους κανόνες της γραμματικής για να τον μετατρέψουμε. (σωστό)
<input type="checkbox"/>	Είναι ήδη σε μορφή CNF. (λάθος)

Τι τύπο γραμματικό πρόβλημα να μετατρέπουν σε CNF;

Απάντηση	
<input checked="" type="checkbox"/> ✖	Τύπου 0 (λάθος)
<input type="checkbox"/>	Τύπου 1 (λάθος)
<input type="checkbox"/>	Τύπου 2 (σωστό)
<input type="checkbox"/>	Τύπου 3 (σωστό)

Οι κανόνες CNF περιλαμβάνουν τον κανόνα $A \rightarrow B$ όπου A και B είναι τελεστές. Οι κανόνες CNF περιλαμβάνουν τον κανόνα $A \rightarrow B$ όπου A και B είναι τελεστές. Οι κανόνες CNF περιλαμβάνουν τον κανόνα $A \rightarrow B$ όπου A και B είναι τελεστές.

Έστω οι παρακάτω κανόνες μετασχηματισμού. Ταράζει καθόλου με τον τύπο της γραμματικής στον οποίο μπορεί να εφαρμοστεί. Επιλέξτε τον πιο περιγραφικό τύπο (το μεγαλύτερο απδεκτό εύρος).

Απάντηση	
<input type="checkbox"/>	$F \rightarrow x \vee F$ Τύπου 0 / Τύπου 3 ✖
<input type="checkbox"/>	$E \vee F \rightarrow T \vee G$ Τύπου 1 / Τύπου 0 ✖
<input type="checkbox"/>	$E \vee F \rightarrow E \vee T$ Τύπου 2 / Τύπου 1 ✖
<input type="checkbox"/>	$G \rightarrow W \vee E$ Τύπου 3 / Τύπου 2 ✖

Ποιοί από τους παρακάτω κανόνες είναι σε μορφή CNF;

Απάντηση	
<input checked="" type="checkbox"/> ✔	$A \leftrightarrow B \vee C$ (σωστό)
<input type="checkbox"/>	$B \leftrightarrow C \vee a$ (λάθος)
<input type="checkbox"/>	$C \leftrightarrow D$ (λάθος)
<input type="checkbox"/>	$D \leftrightarrow b$ (σωστό)
<input type="checkbox"/>	$E \leftrightarrow D \vee A \vee C$ (λάθος)
<input type="checkbox"/>	$F \leftrightarrow a \vee b$ (λάθος)
<input type="checkbox"/>	$G \leftrightarrow F \vee C$ (σωστό)
<input type="checkbox"/>	$H \leftrightarrow a \vee b \vee c$ (λάθος)

Ποιές από τις παρακάτω συμβολοσειρές αναγνωρίζει η ακόλουθη γραμματική;

$S \rightarrow B \vee C$
 $A \rightarrow a \vee c$
 $A \rightarrow S \vee C$
 $B \rightarrow b \vee A$
 $B \rightarrow A$
 $B \rightarrow C \vee a \vee b$
 $C \rightarrow c$

Απάντηση

<input checked="" type="checkbox"/> ✖	abca (λάθος)
<input type="checkbox"/>	bacca (λάθος)
<input type="checkbox"/>	cabc (σωστό)
<input type="checkbox"/>	bacc (σωστό)

Στον αλγόριθμο CKY, για να γρίσουμε το τετράγωνο (1,4) θα πρέπει να συνδυάσουμε τα περιεχόμενα των τετράγωνων:

Απάντηση	
<input checked="" type="checkbox"/> ✖	(1,3) με (2,4) (λίστος)
<input type="checkbox"/>	(1,2) με (3,4) (λίστος)
<input type="checkbox"/>	(1,2) με (2,3) και (2,3) με (3,4) (λίστος)
<input type="checkbox"/>	(1,2) με (3,4) και (1,3) με (2,4) (λίστος)
<input type="checkbox"/>	(1,2) με (2,4) και (1,3) με (3,4) (σωστό)
<input type="checkbox"/>	(1,2) με (2,3) και (1,3) με (2,4) και (2,3) με (3,4) (λίστος)

Στα τετράγωνα του τριγωνικού πίνακα στον αλγόριθμο CKY βάζουμε τερματικά σύμβολα.

Απάντηση	
<input checked="" type="checkbox"/> ✖	Σωστό (λίστος)
<input type="checkbox"/>	Λάθος (σωστό)

Για να χρησιμοποιήσουμε τον αλγόριθμο CKY, η γραμματική τύπου 2 πρέπει να είναι σε μορφή CNF.

Απάντηση	
<input checked="" type="checkbox"/> ✔	Σωστό (σωστό)
<input type="checkbox"/>	Λάθος (λίστος)

--

Στον αλγόριθμο CKY, το τμήμα του πίνακα που βρίσκεται κάτω από την κύρια διαγώνιο παραλείπεται επειδή:

Απάντηση	
<input checked="" type="checkbox"/> ✔	Τα τετράγωνα του δεν απαιτούνται σε σωστά διαστήματα της προς ανάλυση πρότασης. (σωστό)
<input type="checkbox"/>	Αν υπήρχαν, θα περιέχον την ίδια πληροφορία με τον άλλο μισό πίνακα. (λίστος)

Το μέγεθος του πίνακα στον αλγόριθμο CKY εξαρτάται από:

Απάντηση	
<input checked="" type="checkbox"/> ✖	Το πλήθος των κανόνων της γραμματικής. (λίστος)
<input type="checkbox"/>	Το πλήθος των μη-τερματικών συμβόλων. (λίστος)
<input type="checkbox"/>	Το πλήθος των λέξεων της πρότασης που αναλύεται. (σωστό)
<input type="checkbox"/>	Το μέγεθος του λεξικού. (λίστος)
<input type="checkbox"/>	Τίποτα από τα παραπάνω. (λίστος)

Στα τετράγωνα του τριγωνικού πίνακα στον αλγόριθμο CKY δεν ξενοθεύουμε σύμβολα που ήδη υπάρχουν.

Απάντηση	
<input checked="" type="checkbox"/> ✖	Σωστό (λίστος)
<input type="checkbox"/>	Λάθος (σωστό)

--

Έστω το σύνολο κανόνων PCFG:

S → a

[0,3]

S → b

[0,2]

S → b

[0,5]

Η σωστή συγχώνευση είναι:

Απάντηση

☒

✖

S → a

[0,3]

S → b

[0,2]

(λάθος)

☐

S → a

[0,3]

S → b

[0,5]

(λάθος)

☐

S → a

[0,6]

S → b

[0,4]

(λάθος)

☐

S → a

[0,3]

S → b

[0,7]

(σωστό)

Έστω η PCFG και ο πίνακας CKY που ακολουθούν:

S	→	A	C	[0,4]
S	→	C	A	[0,6]
C	→	A	B	[0,2]
C	→	B	A	[0,8]
A	→	a	[0,5]	
A	→	c	[0,3]	
A	→	d	[0,2]	
B	→	b	[0,4]	
B	→	d	[0,6]	

Στο πάνω-δεξιά τετράγωνο του πίνακα εμφανίζονται δύο σύμβολα S, που σημαίνει ότι υπάρχουν δύο τρόποι να αναλυθεί συντακτικά η πρόταση "abc". Αν η πιθανότητα του πρώτου S είναι 0,0192, ποιά είναι η πιθανότητα του δεύτερου;

Απάντηση

☐

✖

0,3908

(λάθος)

☐

0,0384

(λάθος)

☐

0,0072

(σωστό)

☐

0,0192

(λάθος)

Έστω η PCFG:
S → A S [0,2]
S → S A [0,8]
A → a [0,3]
A → S [0,7]

Για να μετατρέψουμε τον τελευταίο κανόνα σε μορφή CNF τον αντικαθιστούμε με τους κανόνες:

Απάντηση

☒

✔

A → A S

[0,14]

A → S A

[0,56]

(σωστό)

☐

A → A S

[0,2]

A → S A

[0,8]

(λάθος)

☐

S → A S

[0,14]

S → S A

[0,56]

(λάθος)

☐

S → A S

[0,2]

S → S A

[0,8]

(λάθος)

Οι πιθανότητες στα τετράγωνα του πίνακα του πιθανοτικού CKY συνάδουν με το γνώσιμο πιθανοτήτων των δύο τετραγώνων συμβόλων στο χώρο των λέξεων κανόνων.

Απάντηση

☐

✖

Σωστό λάθος

☐

Λάθος σωστό

Οι πιθανότητες των τετραγώνων των δύο τετραγώνων συμβόλων στο χώρο των λέξεων και στη πιθανότητα του κανόνα (0,0,0,0) είναι οι αριθμοί που αναγράφονται.

Κανόνες με αυτοαναφορές στις PCFG επιλέγονται και οι πιθανότητες τους μετατρέπονται αναλογικά στους υπάλληλους κανόνες με το ίδιο height.

Απάντηση

☒

✔

Σωστό (σωστό)

☐

Λάθος (λάθος)

Έστω το σύνολο κανόνων PCFG:

S → a

[0,3]

S → b

[0,1]

S → S

[0,6]

Με την απαλοιφή της αυτοαναφοράς το σύνολο γίνεται:

Απάντηση

☒

✖

S → a

[0,3]

S → b

[0,1]

(λάθος)

☐

S → a

[0,75]

S → b

[0,25]

(σωστό)

☐

S → a

[0,18]

S → b

[0,06]

(λάθος)

☐

S → a

[0,6]

S → b

[0,4]

(λάθος)

Η υπόθεση Markon λέει ότι κάθε λέξη στην πρόταση εξορτάται...

Απάντηση

☒ **x**

Σωστό (σωστό)

☐

μόνο από τις προηγούμενες της.
(αυτό)

☐

μόνο από τις επόμενες της.
(αυτό)

☐

και από τις προηγούμενες και από τις επόμενες της.
(αυτό)

☐

από το μοντέλο μονογραμμάτων.
(αυτό)

☐

από το πλήθος εμφανίσεών της στο σώμα κειμένου.
(αυτό)

Στο ίδιο σώμα κειμένου, ο πίνακας εμφανίσεων διγραμμάτων είναι πιο σπαστός από τον πίνακα εμφανίσεων τριγραμμάτων.

Απάντηση

☒ **x**

Σωστό (αυτό)

☐

Λάθος (σωστό)

Τα μοντέλα μονογραμμάτων περιέχουν συντακτική πληροφορία.

Απάντηση

☒ **x**

Σωστό (λάθος)

☐

Λάθος (σωστό)

Τα μηδενικά στους πίνακες πιθανοτήτων των μοντέλων γλώσσας είναι ανεπιθύμητα επιπλέον.

Απάντηση

☒ **x**

σπαταλούν μνήμη στον αποθηκευτικό χώρο του υπολογιστή.
(αυτό)

☐

μηδενίζουν τις πιθανότητες ακόμα και σωστών προτάσεων.
(σωστό)

☐

δεν μεταφέρουν καθόλου πληροφορία.
(αυτό)

☐

δεν μπορούν να υπολογιστούν τα log probs για αυτές τις πιθανότητες.
(αυτό)

Τα μοντέλα διγραμμάτων περιέχουν συντακτική πληροφορία.

Απάντηση

☒ **✓**

Σωστό (σωστό)

☐

Λάθος (λάθος)

Θα θέλατε να σώσετε αυτό το έγγραφο;
<-> ποτέ μη λες ποτέ </>
<-> μη λες ποτέ ποτέ </>
<-> ποτέ ποτέ μη λες </>
Υπολογίστε τις παρακάτω πιθανότητες, ταυροδονώντας με τους σωστούς αριθμούς.

Απάντηση

P(ποτέ | ποτέ)

0.000 / 0.333 **x**

P(λες | μη)

0.000 / 1 **x**

P(μη | λες)

0.000 / 0.000 **✓**

P(ποτέ | λες)

0.000 / 0.667 **x**

P(λες | ποτέ)

0.000 / 0.000 **✓**

Εστω το σήμα:

<-> πρέ μη λες ποτέ </->

<-> πρέ μη λες ποτέ </->

<-> πρέ ποτέ μη λες </->

Υπολόγισε τις πιθανότητες μετά την εφαρμογή εξόμλωσης Laplace, συμπεριλάττες με τους σωστούς αριθμούς.

Απάντηση	
P(λες μ)	6/125 / 0.5 ✖
P(λες λες)	6/273 / 0.125 ✖
P(ποτέ ποτέ)	0.273 / 0.273 ✔
P(μη λες)	0.125 / 0.125 ✔

Τι να απορρίψουμε τις μηδονικές δοσιμεμένες πιθανότητες σε ένα μοντέλο γλώσσας, χρησιμοποιούμε:

(επιλέξτε όλα όσα ταίριαζουν)

Απάντηση	<input checked="" type="checkbox"/> παρεμβολή (interpolation) (σωστό)
	<input type="checkbox"/> μοντέλο ανώτερης τάξης (λάθος)
	<input type="checkbox"/> αλγόριθμο CNN (λάθος)
	<input type="checkbox"/> σπασμόγνωση (backoff) (σωστό)
	<input type="checkbox"/> εξομλωση Laplace (σωστό)
	<input type="checkbox"/> απόσταση Levenshtein. (λάθος)
	<input type="checkbox"/> μετατροπή σε CNE. (λάθος)

Στον πίνακα απεικονίζων τριγράμματα ενός μοντέλου γλώσσας βλέπουμε ένα μηδενικό. Αν εφαρμόσουμε εξομλωση Laplace, αυτό το 0 θα γίνει:

Απάντηση	<input checked="" type="checkbox"/> -1 (λάθος)
	<input type="checkbox"/> 0 (λάθος)
	<input type="checkbox"/> 1 (σωστό)
	<input type="checkbox"/> 2 (λάθος)
	<input type="checkbox"/> 3 (λάθος)
	<input type="checkbox"/> Τίποτα να γράφουμε και τα υπόλοιπα στοιχεία του πίνακα για να απαντήσουμε. (λάθος)

Στα μοντέλα γλώσσας μετάζεις τάζεις (τετραγράμματα, πενταγράμματα, κλπ) ή ντεκα εφραγές το είδος των κεγμάτων του σήματος που χρησιμοποιήσαμε για την κατασκευή τους.

Απάντηση	<input checked="" type="checkbox"/> Σωστό (σωστό)
	<input type="checkbox"/> Λάθος (λάθος)

Στον πίνακα πιθανοτήτων τριγράμματος ενός μοντέλου γλώσσας βλέπουμε ένα μηδενικό. Αν εφαρμόσουμε εξομλωση Laplace, αυτό το 0 θα γίνει:

Απάντηση	<input checked="" type="checkbox"/> -1 (λάθος)
	<input type="checkbox"/> 0 (λάθος)
	<input type="checkbox"/> 1 (λάθος)
	<input type="checkbox"/> 2 (λάθος)
	<input type="checkbox"/> 3 (λάθος)
	<input type="checkbox"/> Δεν έχουμε αρκετές πληροφορίες για να υπολογίσουμε σωφίλως. (σωστό)

Στην τεχνική της σπασμόγνωσης (backoff), όταν η δοσιμεμένη πιθανότητα ενός n-gram είναι 0, τότε...

Απάντηση	<input checked="" type="checkbox"/> προσθέτουμε 1 στον πίνακα εμφάνισης για αυτήν την πιθανότητα. (λάθος)
	<input type="checkbox"/> υπολογίζουμε την πιθανότητα του αμέσως μικρότερης τάξης n-gram. (σωστό)
	<input type="checkbox"/> αγνοούμε τελείως αυτήν την πιθανότητα στον υπολογισμό. (λάθος)
	<input type="checkbox"/> υπολογίζουμε νέα πιθανότητα με βάση τους συντελεστές λ. (λάθος)

Οι συντελεστές λ στην τεχνική της παρεμβολής (interpolation)...
(επιλέξτε όλα όσα ταιριάζουν)

Απάντηση

☒ ✖

έχουν γινόμενο ίσο με 1.
(λάθος)

☐

έχουν άθροισμα ίσο με 1.
(σωστό)

☐

είναι όλοι μικρότεροι (ή ίσοι) από 1.
(σωστό)

☐

ενδέχεται να είναι αρνητικοί.
(λάθος)

Η πληροφορία που αφορά το στρίψιμο ενός κέρματος (κορίνα-γράμματα, ισοπίθανα ενδεχόμενα) είναι:

Απάντηση

☒ ✖

0 bit
(λάθος)

☐

0.5 bit
(λάθος)

☐

1 bit
(σωστό)

☐

2 bit
(λάθος)

Στο ARPA format οι πιθανότητες των n-grams εμφανίζονται με μορφή Logprob.

Απάντηση

☒ ✔

Σιωστό (σωστό)

☐

Λάθος (λάθος)

Η πληροφορία που αφορά τη ρύθμι ενός ζαριού (με 4ξ ισοπίθανα ενδεχόμενα) είναι:

Απάντηση

☒ ✖

0,17 bit
(λάθος)

☐

0,39 bit
(λάθος)

☐

0,5 bit
(λάθος)

☐

1 bit
(λάθος)

☐

2,58 bit
(σωστό)

Η εντροπία...
(επιλέξτε όλα όσα ταιριάζουν)

Απάντηση

☒ ✔

δείχνει πόσο "τακτοποιημένο" ή "ανακατεμένο" είναι ένα σύστημα.
(σωστό)

☐

καθορίζει πόση πληροφορία χρειάζμαστε για να περιγράψουμε πλήρως ένα σύστημα.
(σωστό)

☐

μετράται σε bits.
(σωστό)

☐

υπολογίζεται από πιθανότητες (και λογάριθμους πιθανοτήτων).
(σωστό)

☐

σχετίζεται με την περιπλοκή (replexity).
(σωστό)

☐

εφαρμόζεται μόνο σε γραμματικές τύπου 2.
(λάθος)

Η περιπλοκή (perplexity)...
(επιλέξτε όλα όσα ταιριάζουν)

Απάντηση

☒ σχετίζεται με την εντροπία.
(σωστό)

☐ υπολογίζεται απευθείας από πιθανότητες.
(σωστό)

☐ εκπαίδει τον μέσο παράγοντα διακλάδωσης των κόμβων του γράφου που περιγράφει τη γλώσσα.
(σωστό)

☐ υπολογίζεται από το WordNet.
(λάθος)

To WordNet...
(επιλέξτε όλα όσα ταιριάζουν)

Απάντηση

☒ είναι ένα ηλεκτρονικό λεξικό.
(σωστό)

☐ χρησιμοποιείται για συντακτική ανάλυση.
(λάθος)

☐ έχει μορφή δικτύου με κόμβους και μεταβάσεις.
(σωστό)

☐ χρησιμοποιείται για την εύρεση συνώνυμων λέξεων.
(σωστό)

☐ περιέχει πληροφορίες για τη μορφολογία των λέξεων.
(λάθος)

☐ είναι ένα είδος treebank.
(λάθος)

Ένα αποτελεσματικό εργαλείο για την κατηγοριοποίηση κειμένων είναι:

Απάντηση

☒ ο αλγόριθμος Levenshtein
(λάθος)

☐ ο αλγόριθμος CKY
(λάθος)

☐ οι κοινωνικές εκφράσεις
(λάθος)

☐ ένα πρόγραμμα συμπίεσης αρχείων
(σωστό)

☐ ο κώδικας ASCII
(λάθος)

☐ ένας δεκαεξαδικός συντάκτης
(λάθος)

Οι κυρίες εννοιολογικές συσχετίσεις που κωδικοποιεί το WordNet αφορούν:

Απάντηση

☒ συνωνυμία
(σωστό)

☐ ανωνυμία
(λάθος)

☐ επωνυμία
(λάθος)

☐ μερωνυμία
(σωστό)

☐ τροπωνυμία
(σωστό)

☐ παρωνυμία
(λάθος)

☐ αντινομία
(λάθος)

☐ υπωνυμία
(σωστό)

☐ παρωνομία
(λάθος)

☐ αντωνυμία
(σωστό)

Ταιριάζει τις έννοιες με τις ερμηνείες τους

Απάντηση

συνωνυμία
Λέξεις με συνηθισμένες έννοιες / Λέξεις με παρόμοιες έννοιες ✕
αντωνυμία
ρήματα που τα εναλλάτουμε ελεύθερα με τον εαυτό / Λέξεις με αντιθετικές έννοιες ✕
υπωνυμία
συστασιαστές με σχέση μερος-ολού / ουσιαστικά που το ενα είναι εξειδίκευση του άλλου ✕
μερωνυμία
Λέξεις με παρόμοιες έννοιες / ουσιαστικά με σχέση μέρος προς όλον ✕
τροπωνυμία
συστασιαστές που τα ενα είναι εφάρμογή του άλλου / ρήματα που τα ενα είναι εξειδίκευση του άλλου ✕

Ταυρίστε τις έννοιες	
Απάντηση	
Synset	Γραμματική τύπου 2 / WordNet ✖
Θορυβώδες κανάλι	FSA / Ορθογραφική διάρθωση ✖
Τρίγραμμο	FSA / Μοντέλο γλώσσας ✖
CNF	FSA / Γραμματική τύπου 2 ✖
Παράληλα FSTs	Levenshtein / PC-KIMMO ✖
Βέλτιστο μονοπάτι	PC-KIMMO / Levenshtein ✖
Εντροπία	PC-KIMMO / Perplexity (Περπλοκλή) ✖
Τελική κατάσταση	WordNet / FSA ✖