

1)α)

πωλησεις -> 1,2,3,4

νεοκτιστος ->1,4

διαμερισμα ->1,2,3,4

ξεπερνω->1

προβλεψη->1

παρουσιαζω->2

ανοδος ->2

ιουλιος ->2,3,4

αυξηση->3,4

υπαρχω->4

οι ->1,2

στις->3,4

τις->1

τον->2,3,4

β)i)3,4

ii)(Δεν επιστρέφει κανένα έγγραφο)

2) Σύμφωνα με την συντηρητική προσέγγιση θα εκτελέσουμε πρώτα το (λάμπα OR κουρτίνα) AND (υπολογιστής OR μολύβι) γιατί τα αποτελέσματα αυτών των or είναι τα μικρότερα ( 49 και 55 έναντι 57 χιλιάδες). Στην πραγματικότητα όμως επειδή οι όροι τραπέζι και καρέκλα εμφανίζονται συχνά στο ίδιο έγγραφο, το (τραπέζι OR καρέκλα) θα έχει αρκετά μικρότερο μήκος από το προβλεπόμενο, οπότε η καλύτερη σειρά θα ήταν πρώτα (υπολογιστής OR μολύβι) AND (τραπέζι OR καρέκλα) και μετά το AND (λάμπα OR κουρτίνα).

3)α) Πρώτα θα γίνει το  $X \text{ AND } Y$  και μετά  $\text{AND } Z$  γιατί το  $X$  και  $Y$  έχουν τα μικρότερα μήκη. Το κόστος θα είναι  $100 + 105 + 110 + N = 315 + N$  με  $N$  να είναι η τομή των  $X$  και  $Y$ . Ως συντηρητική προσέγγιση μπορούμε να πάρουμε  $N = 100$  που είναι η μέγιστη τομή οπότε βγαίνει κόστος = 415.

β) Πρώτα θα κάνουμε το  $X \text{ and } Z$  γιατί γνωρίζουμε ότι έχουν 0 τομή για να μην διαβάσουμε το  $Y$ . Το τελικό κόστος θα είναι  $100 + 110 = 210$  και θα επιστρέψουμε το κενό σύνολο.

4)α) 2,4,7

β)4

5)α) Θα γίνουν 11 συγκρίσεις με τις postings lists ως έχουν.

β)Θα γίνουν 3 συγκρίσεις (Με το 4,16 και 47).Επειδή  $\sqrt{p}=4$  κανονικά θα έπρεπε να μετακινούμαστε κατά 4 στην περίπτωση μας όμως αν μετακινούμαστε κατά 4 ο τελευταίος

skip pointer θα δείχνει εκτός της λίστας, συνεπώς ορίζουμε το  $\text{sqrt}(p)=5$  ώστε ο τελευταίος skip pointer να δείχνει εντός της λίστας.

6) Ένας, γιατί το  $\$$  δείχνει το τέλος της λέξης, άρα μόνο μια από τις μεταθέσεις είναι η σωστή.

7)

		c	h	i	c	k	e	n
	0	1	2	3	4	5	6	7
k	1	1	2	3	4	4	5	6
i	2	2	2	2	3	4	5	6
t	3	3	3	3	3	4	5	6
t	4	4	4	4	4	4	5	6
e	5	5	5	5	5	5	4	5
n	6	6	6	6	6	6	5	4

Cost	Operation	Input	output
1	Insert	*	c
1	Replace	k	h
0	Copy	i	i
1	Replace	t	c
1	Replace	t	k
0	Copy	e	e
0	Copy	n	n

8)

query: chord -> ch ho or rd

Term1: or -> or =  $1/4$

Term2: border -> bo or rd de er =  $2/7$

Term3: lord -> lo or rd =  $2/5$

Term4: morbid -> mo or rb bi id =  $1/8$

Term5: sordid -> so or rd di id =  $2/7$

9) Layla ,Lily: L400

Luna ,Liam:L500

10)Anna:A500 ,Hannah:H500

Ate: A300, Eight: E230