

# Object Oriented Programming Course (CEID\_NNY106)

## Εργαστηριακή Άσκηση

### TextAnalyzerApp

### Απαντήσεις σε Συχνές Ερωτήσεις

## 1. Η EA TextAnalyzerApp

Την περιγραφή της EA θα βρείτε στο Git

([https://github.com/ThramboulidisKleanthis/OOP\\_Course\\_Lab](https://github.com/ThramboulidisKleanthis/OOP_Course_Lab)) και στον παρακάτω σύνδεσμο:

<https://eclass.upatras.gr/modules/document/file.php/CEID1434/LabExercises/EA7-TextAnalyzerApp.pdf>

## 2. Ερωτήσεις

2.1 Ερώτηση - Πως πρέπει να ξεκινήσω να δουλεύω την EA;

2.2 Ερώτηση – Είναι απαραίτητο το 1<sup>ο</sup> iteration να διαβάσει και αποθηκεύει το κείμενο στην μνήμη ;

2.3 Ερώτηση – Είναι απαραίτητο να χρησιμοποιήσω HashMap από την αρχή ;

## 3. Απαντήσεις

### 3.1 Απάντηση Ερώτησης 1 (Q1A)

Όπως είπαμε θα πρέπει να ορίσετε και να περιγράψτε το πρώτο iteration ώστε αυτό να μην έχει μεγάλο βαθμό πολυπλοκότητας για εσάς. Υπάρχουν πολλοί τρόποι για να αναθέσετε το έργο του υπολογισμού των στατιστικών στην μηχανή. Σε κάθε περίπτωση όμως θα πρέπει να έχει η μηχανή στη διάθεση της τα δεδομένα εισόδου ώστε να μπορεί να τα επεξεργαστεί.

Καθώς η είσοδος του προγράμματος που θα αναπτύξετε είναι το κείμενο, το οποίο η EA μας ορίζει πως θα είναι σε ένα αρχείο .txt, η πρώτη και βασική ενέργεια που θα πρέπει να κάνει η μηχανή είναι να έχει πρόσβαση στο κείμενο. Θα χρειαστεί, σε όρους προγραμματισμού, να διαβάσει το κείμενο.

Στο διαδίκτυο μπορείτε να βρείτε κώδικα για ανάγνωση αρχείου αρκεί να κάνετε μια αναζήτηση π.χ. java read file. Σας δίνεται ενδεικτικά ένας τέτοιος σύνδεσμος [https://www.w3schools.com/java/java\\_files\\_read.asp](https://www.w3schools.com/java/java_files_read.asp)

Θα παρατηρήσετε πως το πρόγραμμα διαβάσει το κείμενο γραμμή προς γραμμή (nextLine()) και απλά το εκτυπώνει. Αυτό σας επιτρέπει να ελέγξετε αν η μηχανή έχει διαβάσει σωστά το κείμενο σας.

Μπορείτε ακόμη να ζητήσετε από την μηχανή να σας γράψει ένα πρόγραμμα που κάνει αυτή την δουλειά. Στον παρακάτω σύνδεσμο μπορείτε να δείτε μια τέτοια προσπάθεια.

<https://chat.openai.com/share/50f75fe5-6353-4d10-b3a6-804f735de13a>

Η μηχανή εδώ έκανε και δουλειά που δεν της ζητήθηκε (δες ονομα κλάσης). Εσείς θα πάρετε αυτό που θέλετε.

### Μερικά **σχόλια πάνω στον κώδικα του ChatGPT**

#### A) Όνομα κλάσης που αναπαριστά το πρόγραμμα

Προσέξτε πως το ChatGPT, αν και δεν παρακολουθούσε τις διαλέξεις του μαθήματος, δεν έδωσε όνομα Main στην κλάση που περιέχει την main αλλά TextAnalyzer. Θα μου πείτε βέβαια πολλοί από εσάς, και εμείς δεν παρακολουθούμε τις διαλέξεις σας. Δεν θα σχολίαζα το θέμα αν αποδεικνύατε την ώρα του εργαστηρίου πως γνωρίζετε και εφαρμόζετε αυτά που συζητάμε στις διαλέξεις.

Επιπλέον προσέξτε πως έδωσε και λειτουργικότητα που σχετίζεται με το όνομα της κλάσης.

#### B) Η κλάση StringBuilder

Η κλάση αυτή συστήνεται να χρησιμοποιείται αντί της StringBuffer σε περιπτώσεις που τα στιγμιότυπα της χρησιμοποιούνται από ένα μόνο thread καθώς πιθανότατα είναι πιο γρήγορη. Το παρακάτω απόσπασμα από το javaAPI δίνει τα βασικά χαρακτηριστικά της κλάσης.

A mutable sequence of characters. This class provides an API compatible with `StringBuffer`, but with no guarantee of synchronization. This class is designed for use as a drop-in replacement for `StringBuffer` in places where the string buffer was being used by a single thread (as is generally the case). Where possible, it is recommended that this class be used in preference to `StringBuffer` as it will be faster under most implementations. The principal operations on a `StringBuilder` are the `append` and `insert` methods, which are overloaded so as to accept data of any type.

#### Γ) Η κλάση Text

Στον κώδικα του ChatGPT, όπου το κείμενο αναπαρίσταται ως αλφαριθμητικό, εγώ θα όριζα μια κλάση Text στην οποία θα είχα μεταξύ των μεθόδων τις `load(String FName)` και `analyze()` ή εναλλακτικά τις `getLines()` ή/και `getWords()`. Η επιλογή αυτή κάνει τον κώδικα μας πιο αντικειμενοστρεφή.

Τώρα θα πρέπει να αποφασίσετε αν θα επεξεργάζεστε κάθε γραμμή που διαβάσετε και μετά θα προχωράτε στην επόμενη ή θα διαβάσετε όλο το κείμενο στην μνήμη (όπως έκανε

το ChatGPT) και μετά θα το επεξεργαστείτε για να δημιουργήσετε την πληροφορία που σας ζητά η ΕΑ.

Για την ΕΑ αυτή επιλέγουμε **να αποθηκεύσουμε το κείμενο στην μνήμη**. Αυτό σημαίνει πως πρέπει να ορίσουμε πως θα αναπαραστήσουμε στην μνήμη το κείμενο. Ο ορισμός της κατάλληλης δομής αναπαράστασης του κειμένου επηρεάζει και τον τρόπο με τον οποίο θα το επεξεργαστούμε στη συνέχεια. Αν για παράδειγμα το αποθηκεύσουμε σε μία μεταβλητή String η επεξεργασία του θα είναι δύσκολη. Τα ζητούμενα στατιστικά μπορεί να μας βοηθήσουν στην απόφαση μας για το πως θα δομήσουμε το κείμενο στην μνήμη.

### 3.2 Απάντηση Ερώτησης 2 (Q2A)

ΟΧΙ! Αυτό δεν είναι απαραίτητο να υλοποιηθεί στις πρώτες εκδόσεις του προγράμματος σας.

Μπορείτε να βάλετε το κείμενο hardcoded ως String στο πρόγραμμα σας και να δουλέψετε ξεπερνώντας έτσι τα πιθανά προβλήματα που μπορεί να έχετε με την ανάγνωση του από το αρχείο.

Μπορείτε ακόμη αν έχετε αποφασίσει τη δομή με την οποία θα αποθηκεύετε το κείμενο να το βάλετε hardcoded στη δομή αυτή.

Εναλλακτικά μπορείτε να βάλετε το πρόγραμμα να διαβάζει το κείμενο από την βασική είσοδο (System.in) ή ακόμη και να το δέχεται ως command line argument.

Αργότερα θα βρείτε και τη λύση στα προβλήματα που πιθανόν αντιμετωπίσατε με το διάβασμα του αρχείου ακόμη και με την αναπαράστασή του σε μία άλλη μορφή πιο αποδοτική από αυτή του String.

### 3.3 Απάντηση Ερώτησης 3 (Q3A)

ΟΧΙ! Και δεν θα σας το συνιστούσα καθώς δεν έχετε χρησιμοποιήσει μέχρι σήμερα την κλάση αυτή.

Δουλέψτε στην αρχή αποθηκεύοντας τα στατιστικά που θα υπολογίζετε κατά την διαδικασία επεξεργασίας του κειμένου χρησιμοποιώντας κλάσεις της κατηγορίας Collections.

Αργότερα θα σας είναι πιο εύκολο να χρησιμοποιήσετε HashMap για να δώσετε την δομή των αποτελεσμάτων της ΕΑ.