

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**

**ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ**

**«ΕΡΓΑΣΙΑ ΣΤΟ ΜΑΘΗΜΑ ΘΕΩΡΙΑ ΠΛΗΡΟΦΟΡΙΑΣ»**

**Νίκολας Καραχισαριδης**

**Ε19058**

[n.karahisaridis@gmail.com](mailto:n.karahisaridis@gmail.com)

**Φεβρουάριος 2021**

Περιεχόμενα

**Εγκατάσταση Java**

**Οδηγίες για το πρόγραμμα**

**Λειτουργικότητα**

**Ανάλυση των ερωτημάτων**

**Ερώτημα 0**

**Ερώτημα 1**

**Ερώτημα 2**

**Ερώτημα 3**

**Ερώτημα 4**

**Ερώτημα 5**

**Ερώτημα 6**

**Ερώτημα 7**

**Ερώτημα 8**

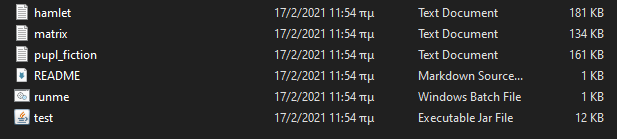
**Εγκατάσταση Java**

**Πατήστε το link :** [Java SE Development Kit 15 - Downloads (oracle.com)](https://www.oracle.com/java/technologies/javase-jdk15-downloads.html)

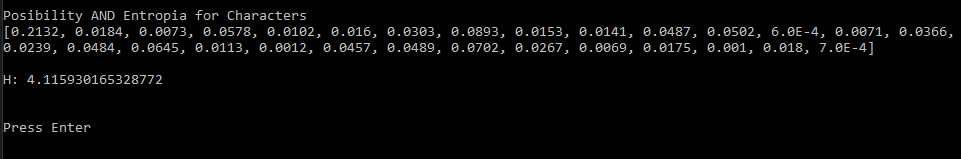
* Και κατεβαστέ για το λογισμικό σας
* Ανοίξτε το αρχείο που κατεβάσατε και τρέξτε το

**Οδηγίες για το πρόγραμμα**

**Ανοίξτε τον φάκελο που έχετε φτιάξει στον C:\ και τρέξτε το αρχείο runme.bat**

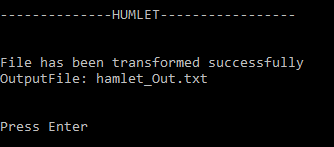


**Θα ανοίξει το CMD και πατώντας το enter θα ολοκληρώνονται οι διαδικασίες μια-μια δίνοντας σας χρόνο να διαβάσετε τα αποτελέσματα.**



**Κάποια δεδομένα λόγο του μεγάλου όγκου αποθηκεύονται σε μορφή αρχείου .**

**(Όταν κάτι αποθηκεύετε σε αρχείο σας αναγράφετε στο cod αντίστοιχο μνήμα και αποθηκεύονται όλα στον φάκελο που φτιάξτε στον C:\ )**



**Λειτουργικότητα**

**Στον φάκελο που θα κάνετε unzip θα βρείτε τα έξεις αρχεία . Τα 3 αρχεία με τα κείμενα-ποιήματα επίσης θα βρείτε το test.jar που είναι ο κώδικας μεταγραμμένες σε runnable αρχείο και ένα αρχείο runme.bat οπού απλά εκκινεί ολύ την διαδικασία και τρέχει τον κώδικα μαζί με κάποιες διαδικασίες για την ομαλή λειτουργεία του κώδικα (κυρίως θέματα ομορφιάς και χωρητικότητας ).**

**Ερώτημα 0**

Αρχικά ανοίγω το αρχείο και περνάω σε μια μεταβλητή όλο το αρχείο (on stream)

Στην συνέχεια με **Ascii.getBytes("US-ASCII")** μετατρέπω σε Ascii χαρακτήρες όλο το κείμενο

Με την βοήθεια κάποιον περιορισμών (If-else) περιορίζω τους Ascii χαρακτήρες σε γράμματα και σύμβολα ΜΟΝΟ (, .) .

Στην συνέχεια με ίδιο τρόπο περιορίζω τα γράμματα σε μικρά (lowercase)

Με την βοήθεια της **.length()** σε ένα String καταφέρνουμε να βρούμε το μέγεθος των χαρακτήρων του κειμένου μας.

Ακριβώς ιδιά διαδικασία και για τα 3 αρχεία (matrix.txt , hamlet.txt , pulp\_fiction.txt)

**Ερώτημα 1**

Αρχικά φτιάχνω έναν πίνακα και με την βοήθεια ενός μετρητή , αυξάνω συνεχώς την αντίστοιχη θέση του πίνακα κατά ένα

Στην συνέχεια αποθηκεύω σε έναν νέο πίνακα τις Πιθανότητες κάνοντας μια διαίρεση το πλήθος του κάθε χαρακτήρα με το μέγεθος του κειμένου .

Με την βοήθεια του παραπάνω πίνακα Πιθανοτήτων βρισκω ευκολά την εντροπία φτιάχνοντας έναν αθροιστή, παίρνοντας την πιθανότητα και τον λογάριθμο του 2 (log2) παίρνοντας μια-μια τις πιθανότητες και αθροίζοντας ΄τες

sumΗ(x) = -pi\*log2pi

Ακριβώς ιδιά διαδικασία και για τα 2 αχρεία (matrix.txt , hamlet.txt.)

**Ερώτημα 2**

Αρχικά μετέτρεψα τους Ascii σε String για να με διευκολύνει στα ζευγάρια με την βοήθεια της String (class ) της οποίας παίρνοντας σαν ορίσματα τον πίνακα μας σε Ascii και το String "UTF-8" μετατρέπομε ξανά τους χαρακτήρες από Ascii σε String .

Με την βοήθεια της .substring() παίρνω τα ζευγαράκια μου ένα-ένα και τα αποθηκεύω σε έναν πίνακα . Ταυτόχρονα σε έναν ακόμα πίνακα καταμετρώ το πλήθος των ζευγαριών.

Έτσι ευκολά υπολογίζουμε την πιθανότητα ,κάνοντας μια διαίρεση το πλήθος του κάθε ζευγαριού με το μέγεθος του πίνακα των ζευγαριών .

Με την βοήθεια του παραπάνω πίνακα Πιθανοτήτων βρισκω ευκολά την εντροπία φτιάχνοντας έναν αθροιστή, παίρνοντας τις αντίστοιχες πιθανότητες και τον λογάριθμο του 2 (log2)με την τομή των δυο πιθανοτήτων παίρνοντας μια-μια τις πιθανότητες και αθροίζοντας ΄τες

sumΗ(x) = -p(I,i+1)\*log2p(I,i+1)

Στην συνέχεια με την βοήθεια ενός ελέγχου σας εκτυπώνω ένα μνήμα αν η από κοινού εντροπία των ζευγαριών είναι μεγαλύτερη από το διπλάσιο της εντροπία των χαρακτήρων.

Για να βρούμε την υπό-συνθήκη εντροπία H(i+1|i) περνούμε την πάρα πάνω εντροπία και την πολλαπλασιάζουμε με τον log2(p(i+1|i))

Ακριβώς ιδιά διαδικασία και για τα 2 αχρεία (matrix.txt , hamlet.txt.)

**Ερώτημα 3**

Επέλεξα η κωδικοποίηση μου να είναι σταθερού μήκους 8 (θα μπορούσε να είναι μήκους 5 , 2ν=29 ν=5). Και το μετέτρεψα σε τους χαρακτήρες μου σε μορφή δυαδική (1,0).

Για τα ζευγάρια έφτιαξα μια διαφορετική κωδικοποίηση πάλι σταθερού μήκους 5 . Ακολουθώντας την εξής λογική: κάθε ζευγάρι αποτελούταν από δυο δυαδικούς αριθμούς με βάση την προηγουμένη κωδικοποίηση έτσι τα μετέτρεψα σε String και τα πρόσθεσα σαν string (concatenate) , και μετα μετέτρεψα αυτόν τον 16ψηφιο σε δεκαδικό

Ακριβώς ιδιά διαδικασία και για τα 2 αχρεία (matrix.txt , hamlet.txt.)

**Ερώτημα 4**

Αρχικά φτιάνω μια κλάση Node με πεδία που θα μπορώ να αποθηκεύσω τον χαρακτήρα , την πιθανότητα του και την κωδικοποιείς του (σε 0,1). Έτσι ,γεμίζω τους πίνακες μου με της πιθανότητες που έχω βρει προηγουμένως και με τους μοναδικούς χαρακτήρες δηλαδή τους χαρακτήρες μου μόνο μια φορά . Με την βοήθεια της κλάσης μπορώ να αντιστοιχήσω κάθε χαρακτήρα με την αντίστοιχη πιθανότητα του. Είμαι λοιπόν έτοιμος να χωρίσω τον πίνακα στην μέση ξανά και ξανά μέχρι την ολοκλήρωση της διαδικασίας και να βάζω μηδενικά στο πρώτο μέρος του πίνακα και άσους στο δεύτερο μέρος του πίνακα . Στο τέλος έχω το επιθυμητό αποτέλεσμα .

Ακριβώς ιδιά διαδικασία και για τα 2 αχρεία (matrix.txt , hamlet.txt.)

**Ερώτημα 5**

**Ακολούθησα την ιδιά λογική με το ερώτημα 4 απλά αυτήν την φορά γέμισα τους πίνακες μου με άλλες πιθανότητες και άλλους χαρακτήρες(ζευγάρια χαρακτήρων) .**

**Ερώτημα 7**

Με την βοήθεια έτοιμης συνάρτησης της java έφτιαξα μια μεταβλητή και την γέμισα με gaussian noise

*java.util.Random r = new java.util.Random();*

*double noise = r.nextGaussian() \* Math.sqrt(9.54242509439) + 0.1;*

*Το 9.54242509439 είναι το 9db σε bps με την βοήθεια του SNR=10log10(S/N)*

Επίσης φτιάχνουμε έναν πίνακα και παίρνοντας χαρακτήρα χαρακτήρα από το encoded αρχείο μας θέτουμε όπου το 0 -> (-1) και όπου 1 -> 1 .

Τέλος προσθέτουμε σε κάθε τιμή του πίνακα την τιμή του gaussian θορύβου που βρήκαμε παραπάνω.

**Ερώτημα 8**

*Ενδέχεται μια μικρή καθυστέρηση για την ολοκλήρωση αυτού του ερωτήματος καθώς είναι αρκετά περίπλοκο και το αρχείο είναι εξίσου μεγάλο .*

Μετατρέπω τον encoding κώδικα σε string για να μπορέσω να πάρω τις 15αδες μου , για να είμαστε σίγουροι προσθέτω στο string στην αρχή κάποια μηδενικά (concatenate) και παίρνω τα τελευταία 15 ψηφιά .Έτσι ώστε η τελευταία μου λέξη να είναι σιγουρά 15 ψηφιά καθώς τα 0νικα στην αρχή δεν αλλάζουν το μνήμα μου (zero-staffing). Στην συνέχεια ορίζω κάποιες θέσεις Εi και κάποιες Πi και ακολουθώ την διαδικασία κωδικοποίησης του hamming.