

ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ 2019 – 2020

ΑΣΚΗΣΗ 1

Νικόλαος Μακρυγεώργος	1115201500238
Αλέξανδρος Ζαφειρίου-Κωστούρος	1115201500040
Παναγιώτης Σταυρόπουλος	1115201500150

Μεταγλώττιση και εκτέλεση :

Η μεταγλώττιση γίνεται μέσω του makefile γράφοντας make όπου παράγεται το εκτελέσιμο με όνομα main

Η εκτέλεση γίνεται με ορίσματα στην γραμμή εντολών τα δύο αρχεία με τις εγγραφές και το μέγεθος κάθε αρχείου σε γραμμές.

Τα ορίσματα μπορούν να δοθούν με οποιαδήποτε σειρά.

π.χ.

```
main -f1 table_R_1000000 -s1 1000000 -f2 table_S_1000000 -s2 1000000
```

Διαγραφή εκτελέσιμου και αντικειμενικών αρχείων γράφοντας make clean

Μετρήσεις Χρόνου:

Για 2 σχέσεις με 1000000 εγγραφές η κάθε μια έχουμε 0,464 seconds.

Για 2 σχέσεις με 2000000 εγγραφές η κάθε μια έχουμε 0,642 seconds.

Για 2 σχέσεις με 5000000 εγγραφές η κάθε μια έχουμε 1,722 seconds.

Για 2 σχέσεις με 10000000 εγγραφές η κάθε μια έχουμε 3,744 seconds.

Περιγραφή Προγράμματος:

main.c :

Αρχικά καλούμε την συνάρτηση take_arguments() η οποία παίρνει τα ορίσματα από την γραμμή εντολών και τα βάζει σε μεταβλητές

Τα ορίσματα είναι τα ονόματα δύο αρχείων τα οποία περιέχουν τις εγγραφές για τις σχέσεις και δύο ακέραιοι που είναι το πλήθος των γραμμών των αρχείων αντίστοιχα.

Στην συνέχεια δημιουργούνται (με malloc) οι πίνακες – σχέσεις και αρχικοποιούνται με τις τιμές από τα αρχεία μέσω της συνάρτησης create_init_relations().

Ακολουθούν δύο κλήσεις της συνάρτησης `recurseFunc()` μια για κάθε σχέση η οποία ταξινομεί τις σχέσεις.

Μετα τις ταξινομήσεις των δυο πινάκων – σχέσεων καλείται η `Sort_Merge_Join()` όπου κάνει το join των σχέσεων και εισάγει τα αποτελέσματα στην λίστα.

Στο τέλος εκτυπώνουμε την λίστα που περιέχει τα αποτελέσματα του join και κάνουμε διαγραφή όλων των πινάκων- σχέσεων και της λίστας.

Επίσης εκτυπώνεται ο `cpu time` και τα αποτελέσματα της λίστα εκτυπώνονται σε αρχείο με όνομα `results.csv`.

Μέσα στην `main` περιέχονται και οι κατάλληλες εντολές για τα tests τα οποία επίσης εκτυπώνονται στο τέλος.

recurseFunc():

Για την ταξινόμηση των σχέσεων δημιουργήσαμε την αναδρομική συνάρτηση `recurseFunc()`.

Παίρνει ορίσματα τον πίνακα – σχέση `R`, τον βοηθητικό πίνακα `R'`, το διάστημα του πίνακα που θα γίνει η ταξινόμηση και έναν ακέραιο που μας λέει με βάση ποια bit θα δημιουργηθεί το ιστόγραμμα.

Εντός της συνάρτησης:

Γίνεται έλεγχος αν το διάστημα που δόθηκε είναι μεγαλύτερο του 4096, αν όχι καλύπτεται η quicksort για το διάστημα αυτό, αν ναι τότε σημαίνει ότι για το διάστημα αυτό θα πρέπει να δημιουργηθεί ο `R'` αρα υπολογίζουμε το οστόγραμμα, το `p_sum` και φτιάχνουμε τον `R'`.

Στην συνέχεια για κάθε ένα bucket που δημιουργήθηκε στον `R'` καλύπτεται η αναδρομική συνάρτηση για κάθε ένα bucket όπου επαναλαμβάνεται η ίδια διαδικασία.

Υπάρχουν σχόλια στον κώδικα όπου είναι απαραίτητο.