

Project (Part 1)

Όνομα : Γιαννούκος Ορέστης

AM : 1115 2016 00035

Όνομα : Λιακόπουλος Γεώργιος

AM : 1115 2016 00091

Όνομα : Χασιαλής Κωνσταντίνος

AM : 1115 2016 00195

Επεξηγήσεις / Παραδοχές στον κώδικα :

Το κομμάτι του sort αποφασίσαμε να το κάνουμε επαναληπτικά και όχι αναδρομικά , καθώς ειπώθηκε στο μάθημα ότι η αναδρομή δεν θα βοηθούσε στην τρίτη φάση του Project. Στην συνάρτηση `iterative_sort` που βρίσκεται στο αρχείο `sort_merge.c` βρίσκεται η υλοποίηση της ταξινόμησης. Επιλέξαμε να επεξεργαζόμαστε τον πίνακα με τα tuples “κατά πλάτος” , δηλαδή σκανάρουμε όλο τον πίνακα φτιάχνουμε το ιστόγραμμα και το `rsum` και κάθε bucket του νέου `reordered` πίνακα το βάζουμε σε μια ουρά . Στην συνέχεια κάθε bucket που κάνουμε pop εξετάζεται αν μπορεί να εφαρμοστεί σε αυτό `quicksort` , αλλιώς ακολουθεί ίδια διαδικασία. Επομένως είναι φανερός ο όρος “κατά πλάτος” αφού στην αρχή θα επεξεργαστούμε πρώτα το 1ο bucket του `reordered` πίνακα , μετά το 2ο κλπ.

Στον έλεγχο για `quicksort` μετράμε τα bytes των tuples + τα bytes του `num_tuples` να είναι μικρότερα από 64 KB.

Στο κομμάτι του `join` διατρέχουμε παράλληλα τους δυο πίνακες με τα tuples και ανάλογα με τις τιμές των κλειδιών αποφασίζουμε το `index` ποιανού πίνακα να αυξήσουμε . Μόλις κάποιο `index` από τα δυο φτάσει να είναι ίσο με `num_tuples` τελειώνει το `join` . Για αυτόν τον λόγο πρέπει να καθορίζεται στο κάλεσμα της συνάρτησης ο πίνακας στο πρώτο όρισμα να έχει το μικρότερο `num_tuples` .

Στο αρχείο `io_functions.c` υπάρχουν διάφορες συναρτήσεις για τον χειρισμό αρχείων που θα δίνονται ως είσοδος . Συγκεκριμένα , μετράμε πρώτα πόσες γραμμές έχει το κάθε αρχείο , δεσμεύουμε την κατάλληλη μνήμη και μετα ξαναδιαβάζουμε το αρχείο για να πάρουμε τα δεδομένα.

Για το `software testing` χρησιμοποιήσαμε την βιβλιοθήκη `Cunit` της C . Τα tests που υποβάλλουμε στο πρόγραμμα μας είναι :

- αν μετράμε σωστά τις γραμμές των αρχείων
- αν διατρέχουμε (`parse`) σωστά το αρχείο
- αν είναι ταξινομημένοι οι πίνακες με τα `relations` μετά την `sort`
- αν γράφτηκαν τα αποτελέσματα επιτυχώς

Για την `join` δεν υπάρχει κάποιο test γιατί πρέπει να ξέρουμε τα αποτελέσματα εξαρχής για να γίνει επαλήθευση.