

Παράλληλος και Κατανεμημένος Υπολογισμός

Εργαστήριο 6

1. Μελετήσετε τον κώδικα 06_Code/StringMatch.

Προσθέστε εντολές μέτρησης χρόνου.

Χρησιμοποιείτε το αρχείο world192.txt από το Large Corpus στο σύνδεσμο <https://corpus.canterbury.ac.nz/descriptions/#large> (περίπου 2M)

Μπορείτε για παράδειγμα να τετραπλασιάσετε το μέγεθος του αρχείου ως εξής:

Windows: `type world192.txt world192.txt world192.txt world192.txt > fourtimes.txt`

Linux: `cat world192.txt world192.txt world192.txt world192.txt > fourtimes.txt`

Διαλέξτε ένα pattern, για παράδειγμα “Exclusive economic zone”

Μετρήστε το χρόνο εκτέλεσης.

Χρησιμοποιείτε τις τεχνικές απεικόνισης και αναγωγής για να παραλληλίσετε τον κώδικα.

Μετρήστε το χρόνο εκτέλεσης.

Προσοχή στην κατανομή των blocks: κατανέμουμε τον πίνακα των αποτελεσμάτων, όχι το κείμενο.

Στο εσωτερικό βρόχο το κείμενο επικαλύπτεται κατά $m-1$ χαρακτήρες με το επόμενο block του κειμένου (όπου m το μέγεθος του pattern) ώστε να καλυφθεί η πιθανότητα ένα pattern να βρίσκεται στα σύνορα δυο γειτονικών blocks.

Προσοχή στην άθροιση του αριθμού των ταυτίσεων: εφαρμόστε τις τεχνικές αναγωγής.

2. Μελετήσετε τον κώδικα 06_Code/StringHistogram.

Προσθέστε εντολές μέτρησης χρόνου.

Χρησιμοποιείτε το ίδιο αρχείο όπως στην άσκηση 1.

Μετρήστε το χρόνο εκτέλεσης.

Χρησιμοποιείτε τις τεχνικές απεικόνισης και αναγωγής για να παραλληλίσετε τον κώδικα.

Μετρήστε το χρόνο εκτέλεσης.

Προσοχή στον τρόπο συλλογής των συχνοτήτων: έχουμε αναγωγή σε δομή δεδομένων (πίνακα) όχι απλό βαθμωτό μέγεθος (μια μεταβλητή).

Αν δημιουργήσετε μόνο ένα μοιραζόμενο πίνακα πόσες φορές πρέπει να εκτελέσετε αμοιβαίο αποκλεισμό; Χρησιμοποιήστε τοπικές δομές για μείωση των συγκρούσεων.

Αν εφαρμόσετε αμοιβαίο αποκλεισμό στο σύνολο του πίνακα ποιο θα είναι το αποτέλεσμα στο χρόνο εκτέλεσης;

3. Μελετήσετε τον κώδικα 06_Code/WordCount.

Προσθέστε εντολές μέτρησης χρόνου.

Χρησιμοποιείτε το ίδιο αρχείο όπως στην άσκηση 1.

Δοκιμάστε με τις δομές TreeMap και HashMap.

Η δομή TreeMap παράγει ταξινομημένο αποτέλεσμα ενώ η HashMap όχι. Ανατρέξτε στη υλοποίηση του TreeMap για να αντιληφθείτε τη διαφορά.

Μετρήστε το χρόνο εκτέλεσης.

Χρησιμοποιείτε τις τεχνικές απεικόνισης και αναγωγής για να παραλληλίσετε τον κώδικα.

Δοκιμάστε με τις δομές TreeMap και HashMap.

Χρησιμοποιείτε τοπικές δομές ανά νήμα για να μειώσετε τις συγκρούσεις του αμοιβαίου αποκλεισμού.

Ποιά είναι η διαφορά στο παραλληλισμό των δύο δομών; Ποιά είναι πιο συμφέρουσα στο παραλληλισμό; Γιατί κατά τη γνώμη σας η Java δεν παρέχει ConcurrentTreeMap;

Μετρήστε το χρόνο εκτέλεσης.

Σημείωση:

Οι κώδικες είναι οι απλοί. Υπάρχουν πιο αποδοτικές και κομψές υλοποιήσεις, που όμως καλούν συναρτήσεις βιβλιοθήκης που δεν είναι εύκολο να ελέγξουμε.