

ΠΑΡΑΛΛΗΛΟΣ ΥΠΟΛΟΓΙΣΜΟΣ

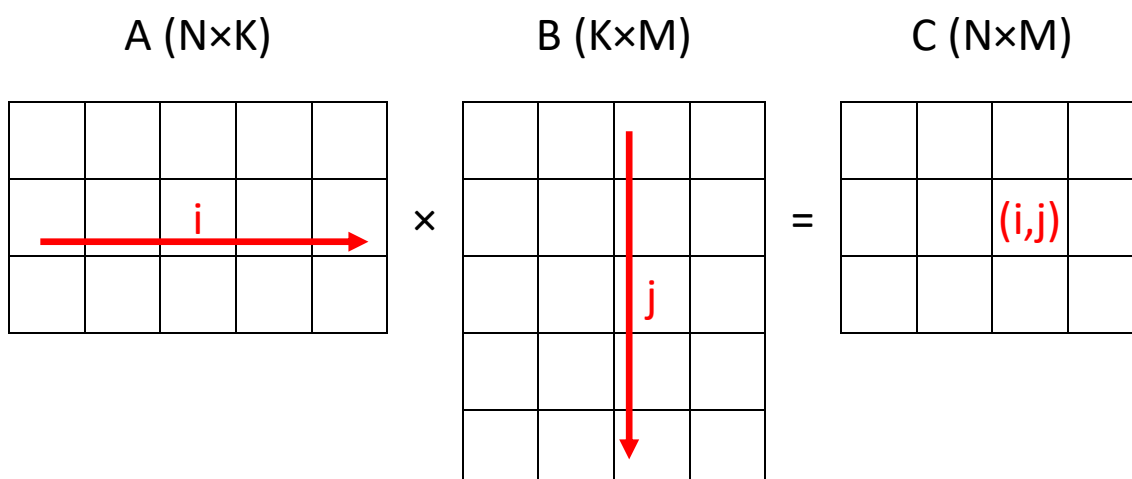
ΑΚΑΔΗΜΑΪΚΟ ΕΤΟΣ 2022/2023

Εισαγωγή

Στα πλαίσια της άσκησης θα ασχοληθούμε με έναν απλό και ευρέως γνωστό υπολογισμό: τον πολλαπλασιασμό δύο 2Δ μητρώων. Θα πρέπει κατ' αρχήν να μελετήσετε την επίδοση του ακολουθιακού αλγόριθμου και στην συνέχεια να υλοποιήσετε παράλληλους αλγορίθμους για τον υπολογισμό αυτό χρησιμοποιώντας POSIX Threads και OpenMP.

Πολλαπλασιασμός 2Δ μητρώων

Το αποτέλεσμα του πολλαπλασιασμού ενός 2Δ μητρώου A μεγέθους $N \times K$ με ένα 2Δ μητρώο B μεγέθους $K \times M$ είναι ένα 2Δ μητρώο C μεγέθους $N \times M$. Κάθε γραμμή του μητρώου A πρέπει να πολλαπλασιαστεί με κάθε στήλη του μητρώου B. Ο πολλαπλασιασμός της γραμμής i του μητρώου A με την στήλη j του μητρώου B δίνει το στοιχείο $C(i, j)$ (δείτε και το παρακάτω σχήμα).



Θεωρώντας πως όλα τα στοιχεία του μητρώου C έχουν αρχικοποιηθεί σε 0, ο αλγόριθμος που εκφράζει άμεσα την παραπάνω περιγραφή είναι:

```
for i = 0 to N - 1
  for j = 0 to M - 1
    for k = 0 to K - 1
      C(i, j) += A(i, k) * B(k, j)
    endfor
  endfor
endfor
```

Όπως όμως είναι εύκολα αντιληπτό, ο υπολογισμός κάθε στοιχείου $C(i, j)$ δεν εξαρτάται από άλλα στοιχεία του μητρώου C. Επομένως, κάθε μετάθεση των τριών βρόχων είναι επιτρεπτή και δίνει το σωστό τελικό αποτέλεσμα. Είναι εύκολο να δούμε πως το πλήθος των μεταθέσεων είναι $3! = 1 \cdot 2 \cdot 3 = 6$.

Ζητούμενα της άσκησης

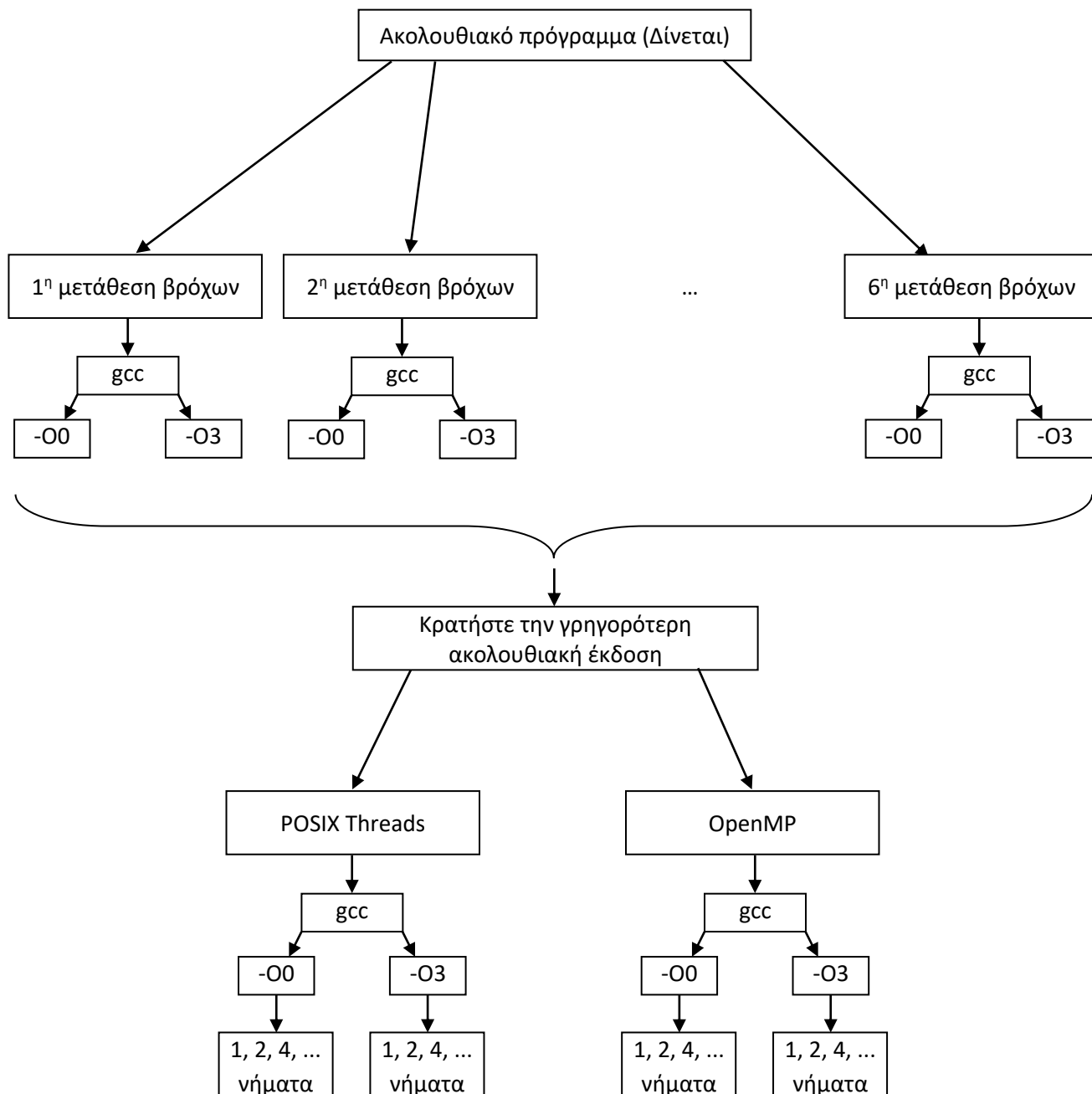
Στα πλαίσια της άσκησης θα σας δοθεί μια ακολουθιακή υλοποίηση του πολλαπλασιασμού μητρώων. Βασιζόμενοι σε αυτήν, θα πρέπει κατ' αρχήν να κάνετε μια μελέτη της επίδοσης της υλοποίησης αυτής και στην συνέχεια να φτιάξετε δύο παράλληλες υλοποιήσεις (μία με χρήση POSIX Threads και μία με χρήση OpenMP). Συγκεκριμένα:

- 1) **Πριν ξεκινήσετε την παραλληλοποίηση του κώδικα**, εκτελέστε τον ακολουθιακό κώδικα για τους παρακάτω συνδυασμούς:
 - a. Τις 6 πιθανές μεταθέσεις των 3 βρόχων.
 - b. Μεταγλώττιση του προγράμματος χωρίς βελτιστοποίηση (gcc -O0) και με μέγιστη βελτιστοποίηση (gcc -O3).
 - c. Διαφορετικά μεγέθη μητρώων (δείτε το Παράρτημα Β)Δημιουργήστε γραφήματα για τους χρόνους εκτέλεσης (ομαδοποιώντας κατάλληλα). Καταγράψτε τις παρατηρήσεις σας στην αναφορά που θα παραδώσετε και προσπαθήστε να αιτιολογήσετε τις διαφορές στους χρόνους εκτέλεσης.
- 2) Για τον ταχύτερο συνδυασμό μετάθεσης βρόχων και επιπέδου βελτιστοποίησης, δημιουργήστε παράλληλα προγράμματα για τον πολλαπλασιασμό των μητρώων χρησιμοποιώντας POSIX Threads και OpenMP. Εκτελέστε τις παράλληλες εκδόσεις των προγραμμάτων σας για τα ίδια μητρώα που χρησιμοποιήσατε και στην ακολουθιακή έκδοση. Για κάθε μητρώο εκτελέστε το πρόγραμμα με διαφορετικό πλήθος νημάτων και δημιουργήστε γραφήματα της **χρονοβελτίωσης (speedup)** του προγράμματος σας. Καταγράψτε τις παρατηρήσεις σας στην αναφορά που θα παραδώσετε.
- 3) Για την έκδοση του παράλληλου προγράμματος σας με OpenMP δοκιμάστε διαφορετικές προσεγγίσεις για τον διαμοιρασμό των επαναλήψεων των βρόχων σε νήματα (collapse, static, dynamic, guided, επίπεδο εμφωλευμένου βρόχου όπου γίνεται η παραλληλοποίηση,...). Δικαιολογήστε τυχόν διαφορές στον χρόνο εκτέλεσης.

Εκτός όλων των παραπάνω είσαστε προφανώς ελεύθεροι να βελτιστοποιήσετε και να παραλληλοποιήσετε τον κώδικα με όποιον άλλο τρόπο θέλετε. Αν θεωρήσετε πως μια οποιαδήποτε αναδιοργάνωση του κώδικα μπορεί να οδηγήσει σε καλύτερη παραλληλοποίηση και απόδοση είστε ελεύθεροι να το κάνετε. Ωστόσο, θεωρείται αυτονόητο πως η υλοποίηση θα πρέπει να είναι σωστή (να δίνει **πάντα** τα ίδια αποτελέσματα με την ακολουθιακή έκδοση του προγράμματος) και πως **στην αναφορά που θα παραδώσετε θα πρέπει να αιτιολογήσετε γιατί παραλληλοποιήσατε με τον συγκεκριμένο τρόπο την εφαρμογή**.

Οδηγίες για την μεταγλώττιση και εκτέλεση του ακολουθιακού κώδικα που σας δίνεται και του παράλληλου κώδικα που θα φτιάξετε δίνονται στο Παράρτημα Β της εκφώνησης. Εκεί επίσης αναφέρεται για ποιες παραμέτρους θα πρέπει να τρέξετε την κάθε έκδοση του προγράμματος και ποια αποτελέσματα θα συμπεριλάβετε στην αναφορά σας.

Συγκεντρωτικά, όλες οι περιπτώσεις που θα πρέπει να συμπεριλάβετε φαίνονται στο παρακάτω σχήμα:



Παραδοτέα

Τα παραδοτέα για την εργασία σας είναι μια γραπτή αναφορά και ο κώδικας της άσκησης που θα αναπτύξετε. **Η προθεσμία παράδοσης της εργασίας ορίζεται η Τρίτη 02/06/2023 και ώρα 23:59:59. Η εργασία θα πρέπει να παραδοθεί αποκλειστικά μέσω της ηλεκτρονικής πλατφόρμας GUNET2 (εργασίες που θα αποσταλούν μέσω e-mail δεν θα βαθμολογηθούν) και η πλατφόρμα θα κλειδώσει αυτόματα την δυνατότητα υποβολής εργασιών μετά την λήξη της προθεσμίας. Οργανώστε λοιπόν σωστά τον χρόνο σας. Μετά την είσοδο σας στο σύστημα θα πρέπει να μεταβείτε στο μάθημα “Παράλληλος Υπολογισμός” και στο μενού αριστερά να μεταβείτε στο “Εργασίες”. Κάθε ομάδα θα παραδώσει μια φορά μόνο την εργασία (όχι κάθε φοιτητής ξεχωριστά).**

Στην αναφορά **δεν** θα πρέπει να περιλαμβάνεται επεξήγηση του ακολουθιακού αλγόριθμου. Επικεντρωθείτε στην επεξήγηση της παραλληλοποίησης που κάνατε, στις μετρήσεις σας και στα διαγράμματα που θα προσθέσετε. Σχολιάστε τις διαφορές από την χρήση βελτιστοποιήσεων στον χρόνο εκτέλεσης της εφαρμογής και (κυρίως) στην χρονοβελτίωση. Γενικότερα, δώστε ιδιαίτερο βάρος στην αναφορά σε αυτά που κάνατε εσείς.

Ο βαθμός της εργασίας αποτελεί το 30% της τελικής βαθμολογίας. Το υπόλοιπο 70% προκύπτει από την τελική εξέταση. Για να περάσει κάποιος φοιτητής το μάθημα δεν είναι απαραίτητη η παράδοση της εργασίας. **Στην περίπτωση αυτή ωστόσο, θεωρείται πως η εργασία έχει πάρει βαθμό 0 (μηδέν). Ο τελικός βαθμός τότε προκύπτει μόνο από το 70% της τελικής εξέτασης και θα πρέπει να είναι προβιβάσιμος (≥ 5).**

Ο βαθμός της εργασίας διατηρείται μέχρι και την άτυπη εξεταστική Φεβρουαρίου 2024. Αν κάποιος φοιτητής δεν περάσει το μάθημα μέχρι τότε και αποφασίσει να το παρακολουθήσει σε επόμενο ακαδημαϊκό έτος θα πρέπει να ανταποκριθεί στις υποχρεώσεις του μαθήματος για το ακαδημαϊκό έτος που θα το παρακολουθήσει ξανά.

Σημειώνεται επίσης πως δεν είναι δυνατή η παράδοση της εργασίας σε άλλη ημερομηνία, π.χ. κατά την εξεταστική του Σεπτεμβρίου ή την άτυπη εξεταστική Φεβρουαρίου.

Παράρτημα Α

Για την εγκατάσταση των προγραμμάτων που θα χρειαστείτε στα πλαίσια της άσκησης, θα πρέπει να ακολουθήσετε τις παρακάτω οδηγίες.

1) Εγκατάσταση Linux

Στα πλαίσια της εργασίας συστήνεται να εγκαταστήσετε την έκδοση 20.04.2 LTS της διανομής Ubuntu (<http://www.ubuntu.com>). Οι οδηγίες που ακολουθούν υποθέτουν την εγκατάσταση της συγκεκριμένης διανομής. Φυσικά μπορείτε να εγκαταστήσετε όποια άλλη διανομή θέλετε, όπως για παράδειγμα Fedora (<http://fedoraproject.org>) ή openSUSE (<http://www.opensuse.org>). Σε κάθε περίπτωση επιλέξτε την τελευταία έκδοση που είναι διαθέσιμη για κάθε διανομή. Αν εγκαταστήσετε άλλη διανομή εκτός της προτεινόμενης θα πρέπει να προσαρμόσετε κατάλληλα της οδηγίες για την εγκατάσταση των επιπλέον πακέτων που απαιτούνται.

Οι διανομές συνήθως είναι διαθέσιμες ως αρχεία τύπου ISO. Αν εγκαταστήσετε την διανομή απευθείας στον σκληρό δίσκο του υπολογιστή σας (βλέπε παρακάτω) θα πρέπει να γράψετε το αρχείο ISO σε ένα CD, DVD ή USB stick. Αν εγκαταστήσετε την διανομή σε ένα Virtual Machine (βλέπε παρακάτω) μπορείτε να χρησιμοποιήσετε απευθείας το αρχείο ISO. Η εγκατάσταση της διανομής μπορεί να γίνει με δύο τρόπους:

a. Απευθείας στον σκληρό δίσκο του μηχανήματος σας

Αν ακολουθήσετε την μέθοδο αυτή θα πρέπει να φτιάξετε ένα ξεχωριστό partition στον σκληρό σας δίσκο. Όταν ξεκινήσετε την εγκατάσταση της διανομής θα πρέπει να επιλέξετε το συγκεκριμένο partition. Ακολουθήστε αυτή τη μέθοδο αν θέλετε να κρατήσετε την διανομή που θα εγκαταστήσετε και για αργότερα.

b. Μέσω Virtual Machine (VM)

“Virtual Machine” είναι στην πραγματικότητα οποιοδήποτε πρόγραμμα που μπορούμε να εγκαταστήσουμε σε ένα λειτουργικό σύστημα που διαθέτουμε και προσομοιώνει έναν υπολογιστή. Αν για παράδειγμα διαθέτουμε Windows, μπορούμε να εγκαταστήσουμε ένα Virtual Machine και στην συνέχεια να εγκαταστήσουμε στον προσομοιούμενο υπολογιστή ένα άλλο λειτουργικό σύστημα, όπως για παράδειγμα Linux.

Αν ακολουθήσετε αυτή τη μέθοδο θα πρέπει κατ’ αρχάς να κατεβάσετε και να εγκαταστήσετε στο λειτουργικό σύστημα που διαθέτετε ένα Virtual Machine. Υπάρχουν αρκετά ελεύθερα προγράμματα διαθέσιμα για αυτό, όπως το Virtual Box (<https://www.virtualbox.org>), **το οποίο και συστήνουμε για την εργασία**, το VMware Player (<http://www.vmware.com/products/player/overview.html>) και το QEMU (<http://wiki.qemu.org>). Επειδή θα χρειαστεί να τρέξουμε παράλληλα προγράμματα, διαβάστε τις οδηγίες εγκατάστασης και δημιουργίας Virtual Machine, ώστε η Virtual Machine στην οποία θα εγκαταστήσετε την διανομή Linux που θα επιλέξετε να υποστηρίξει πολλαπλούς πυρήνες. Επίσης, δηλώστε έναν αρκετά μεγάλο σκληρό δίσκο (τουλάχιστον 10GB). Τέλος, εγκαταστήστε την διανομή Linux που επιλέξατε στην Virtual Machine που φτιάξατε.

2) Εγκατάσταση Guest Additions

Αν έχετε εγκαταστήσει VirtualBox τότε μια σημαντική προσθήκη που μπορείτε να κάνετε είναι τα “Guest Additions”. Τα εργαλεία αυτά προσφέρουν επιπλέον δυνατότητες στον εικονικό υπολογιστή, με κυριότερες την καλύτερη ανάλυση οθόνης και το Clipboard για την αντιγραφή δεδομένων και αρχείων μεταξύ του πραγματικού και του εικονικού υπολογιστή. **Η εγκατάσταση των “Guest Additions” είναι εντελώς προαιρετική. Μπορείτε να ολοκληρώσετε την εργασία και χωρίς αυτά.** Αν θέλετε να τα εγκαταστήσετε ακολουθήστε τα παρακάτω βήματα:

- a. Πριν εκκινήσετε την VM εγκαταστήστε στο VirtualBox το Extension Pack:
<https://www.virtualbox.org/manual/ch01.html#intro-installing>
- b. Εκκινήστε την VM και εγκαταστήστε τα προαπαιτούμενα πακέτα στο λειτουργικό σύστημα του εικονικού υπολογιστή. Από την γραμμή εντολών εκτελέστε την εντολή:

```
sudo apt-get install dkms
```
- c. Κάντε επανεκκίνηση του εξομοιούμενου υπολογιστή σας.
- d. Στο μενού του Virtual Box επιλέξτε “Devices → Insert Guest Additions CD image...”.
- e. Επιτρέψτε την αυτόματη εκτέλεση του προγράμματος εγκατάστασης.
- f. Κάντε επανεκκίνηση του εξομοιούμενου υπολογιστή σας.
- g. Ανοίξτε το πρόγραμμα Διαχείρισης Αρχείων (File Manager).
- h. Κάντε unmount το “Guest Additions CD” κάνοντας κλικ στο σύμβολο “Λ”.

Παράρτημα Β

- 1) Για την μεταγλώττιση του ακολουθιακού προγράμματος που σας δίνεται μπορείτε να χρησιμοποιήσετε την παρακάτω εντολή:

```
gcc -O3 (ή -O0) -Wall -Wextra -o mm mm.c
```

Αυτή θα δημιουργήσει ένα εκτελέσιμο αρχείο με το όνομα "mm", το οποίο μπορείτε να εκτελέσετε με την εντολή:

```
./mm <Παράμετροι γραμμής εντολής>
```

Οι παράμετροι γραμμής εντολής είναι με την σειρά τα μεγέθη N, K και M των 2Δ μητρώων προς πολλαπλασιασμό, ένας ακέραιος αριθμός αν θα εκτυπωθούν τα αποτελέσματα στην οθόνη ή όχι (τιμή διάφορη του 0 ή 0 αντίστοιχα) και το πλήθος των νημάτων που θα δημιουργηθούν στο πρόγραμμα σας. Η τελευταία παράμετρος δεν χρησιμοποιείται στο ακολουθιακό πρόγραμμα. Έχει προστεθεί προς διευκόλυνση σας για την υλοποίηση με POSIX Threads, αλλά πρέπει να δίνεται στην γραμμή εντολών σε κάθε περίπτωση.

Αν θέλετε να εκτελέσετε το πρόγραμμα για τα μητρώα A και B μεγέθους 100x400 και 400x600 αντίστοιχα και να μην εμφανίσετε τα αποτελέσματα στην οθόνη θα πρέπει να εκτελέσετε το πρόγραμμα ως εξής:

```
./mm 100 400 600 0 1
```

Αν θέλετε να εμφανίσετε τα αποτελέσματα στην οθόνη, αλλάξτε την προτελευταία παράμετρο σε μια οποιαδήποτε μη μηδενική τιμή.

- 2) Εκτελέστε το σειριακό και τα παράλληλα προγράμματα σας για 3 είδη μητρώων:
 - a. Τα τρία μητρώα να είναι τετραγωνικά και να έχουν το ίδιο μέγεθος, π.χ. $N = K = M = 2000$.
 - b. Το μητρώο A να έχει σχετικά λίγες γραμμές και πολλές στήλες, ενώ το B το ανάποδο, π.χ. $N = 100, K = 4000, M = 20000$.
 - c. Το μητρώο A να έχει πολλές γραμμές και σχετικά λίγες στήλες, ενώ το B το ανάποδο, π.χ. $N = 4000, K = 100, M = 20000$.

Οι παραπάνω διαστάσεις των μητρώων θα πρέπει να τροποποιηθούν από εσάς κατάλληλα. Κατά τις δοκιμές σας μπορείτε να χρησιμοποιείτε σχετικά μικρά μητρώα για να επιβεβαιώνετε εύκολα και γρήγορα τα αποτελέσματα σας. Για την αναφορά σας επιλέξτε μεγαλύτερα μεγέθη μητρώων, ώστε οι χρόνοι εκτέλεσης να είναι αρκετά μεγάλοι (> 1 λεπτό για την ακολουθιακή εκτέλεση). Αυτό θα σας επιτρέψει να μελετήσετε με μεγαλύτερη ακρίβεια την επίδραση της χρήσης περισσότερων επεξεργαστών.

Συνίσταται να τρέξετε την ακολουθιακή εφαρμογή με τα σύνολα παραμέτρων που θα επιλέξετε και να αποθηκεύσετε τα αποτελέσματα σε αρχεία, ώστε να έχετε τα αποτελέσματα αυτά σαν αναφορά για τις αλλαγές που θα κάνετε κατά την παραλληλοποίηση. Μην ξεχνάτε πως θα πρέπει να παίρνετε τα ίδια αποτελέσματα από το ακολουθιακό και το παράλληλο πρόγραμμα!

- 3) Για την μεταγλώττιση του παράλληλου προγράμματος με POSIX Threads που θα φτιάξετε μπορείτε να χρησιμοποιήσετε την παρακάτω εντολή (θεωρώντας πως το όνομα του αρχείου που φτιάξατε είναι "mm_pthreads.c"):

```
gcc -O3 (ή -O0) -pthread -Wall -Wextra -o mm_threads  
mm_threads.c
```

Αυτή θα δημιουργήσει ένα εκτελέσιμο αρχείο με το όνομα “mm_threads”. Μπορείτε στην συνέχεια να το εκτελέσετε με την εντολή:

```
./mm_threads <Παράμετροι γραμμής εντολής>
```

- 4) Για την μεταγλώττιση του παράλληλου προγράμματος με OpenMP που θα φτιάξετε μπορείτε να χρησιμοποιήσετε την παρακάτω εντολή (θεωρώντας πως το όνομα του αρχείου που φτιάξατε είναι “mm_omp.c”):

```
gcc -O3 (ή -O0) -fopenmp -Wall -Wextra -o mm_omp mm_omp.c
```

Αυτή θα δημιουργήσει ένα εκτελέσιμο αρχείο με το όνομα “mm_omp”. Μπορείτε να ορίσετε το πλήθος των νημάτων που θα δημιουργούνται σε κάθε παράλληλη περιοχή και στην συνέχεια να το εκτελέσετε με τις εντολές:

```
export OMP_NUM_THREADS=<Πλήθος νημάτων ανά παράλληλη περιοχή>
```

```
./mm_omp <Παράμετροι γραμμής εντολής>
```

- 5) Για να κάνετε εύκολα την σύγκριση αποτελεσμάτων που έχετε αποθηκεύσει σε αρχεία μπορείτε να χρησιμοποιήσετε το πρόγραμμα numdiff. Η ακρίβεια των αποτελεσμάτων πρέπει να είναι τουλάχιστον μέχρι το 12 δεκαδικό ψηφίο:

```
numdiff -a 1e-12 result1.txt result2.txt
```