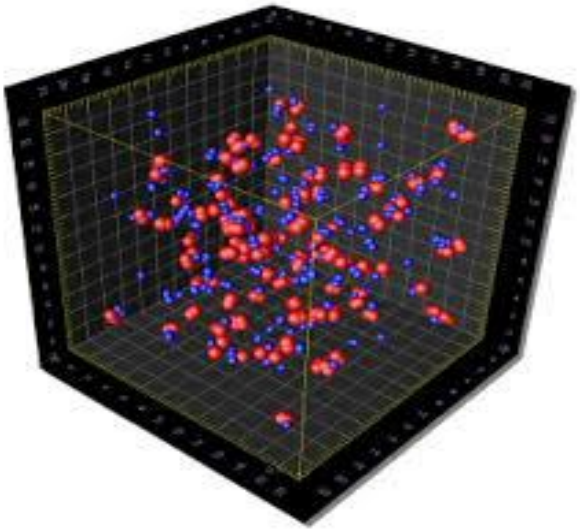
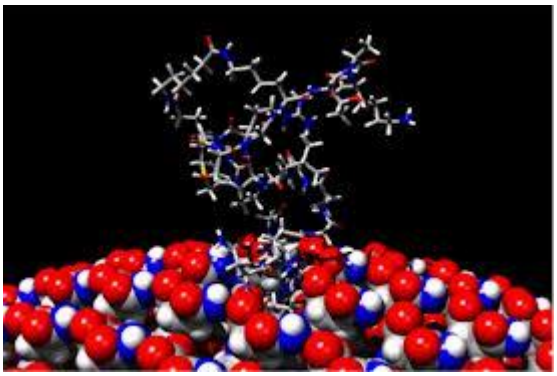


Molecular Dynamics Simulation on GPU

סימולציה של דינמיקה מולקולארית על גבי כרטיס גרפי



סטודנט: ולדימיר נוביקוב

מנחה: דר' חסין יהודה

ת.ז: 312669112

חתימה:

חתימה:

תאריך: 14.02.2014

תוכן עניינים

1. תקציר ..... עמוד 3

2. מסגרת הפרויקט ..... עמוד 3-5

2.א. רקע לפרויקט ..... עמוד 3

2. ב. הסבר על הניסוי המדעי ..... עמוד 4

2. ג. הסבר על הסימולציה הממוחשבת ..... עמוד 4-5

3. הבעיות והפתרונות ..... עמוד 5-8

3.א. בעיית זמני הריצה 5-7

3.ב. בעייתי שגיאות החישוב 7-8

4. אב טיפוס שמומש 8-10

4.א. תיאור אב טיפוס 8

4.ב. תרשים מודולים 9

4.ג. תרשים מחלקות 10

5. בדיקות של המערכת 11-13

6. דרישות טכניות 14

7. סיכונים להמשך 14

8. לוח זמנים להמשך וספרות מקצועית 15

נספחים 16

נספחים א': תכנות מקבילי ב CUDA ..... עמוד 16

1. תקציר

דינמיקה מולקולרית היא סימולציה ממוחשבת המדמה תנועת אטומים ומולקולות של מערכות כימיות פיסיקליות הנגרמת מהקשרים השונים בין האטומים במערכת ושאר הכוחות הפועלים עליהם. בעזרת דינמיקה מולקולארית ניתן לחסוך זמן יקר על חישובים ידניים, לראות את מצב המערכת הכימית המתוארת בכל נקודת זמן נתונה ובכך לאשש או להפריך השערות כימיות ופיסיקליות.

סימולציות כאלה שפועלות במשך זמן ממושך סובלות מחוסר דיוק בתוצאות עקב קיזוזי מספרים במחשב וגדילת גודל השגיאה עם הזמן וכמו כן מזמני ריצה גדולים בגלל ריבוי האטומים המתוארים במערכת.

במסמך זה מתוארת מערכת כימית שבשבילה נכתבת הסימולציה, בעיית השגיאה המספרית, בעיית זמני הריצה ודרכי הפתרון.

בנוסף מצורף לוח הזמנים וסיכונים להמשך עבודה, תיאור אב-טיפוס הסימולציה, בדיקות ונספחים נוספים.

2. מסגרת הפרויקט

הפרויקט נכתב במסגרת פרויקט גמר מחקרי במסלול להנדסת תוכנה במכללת "עזריאלי – המכללה האקדמית להנדסה, ירושלים" תחת הנחיתו של דר' חסין יהודה ובשיתוף פעולה עם פרופ' תמר רז נחום וצוות חוקרים מהאוניברסיטה העברית בירושלים.

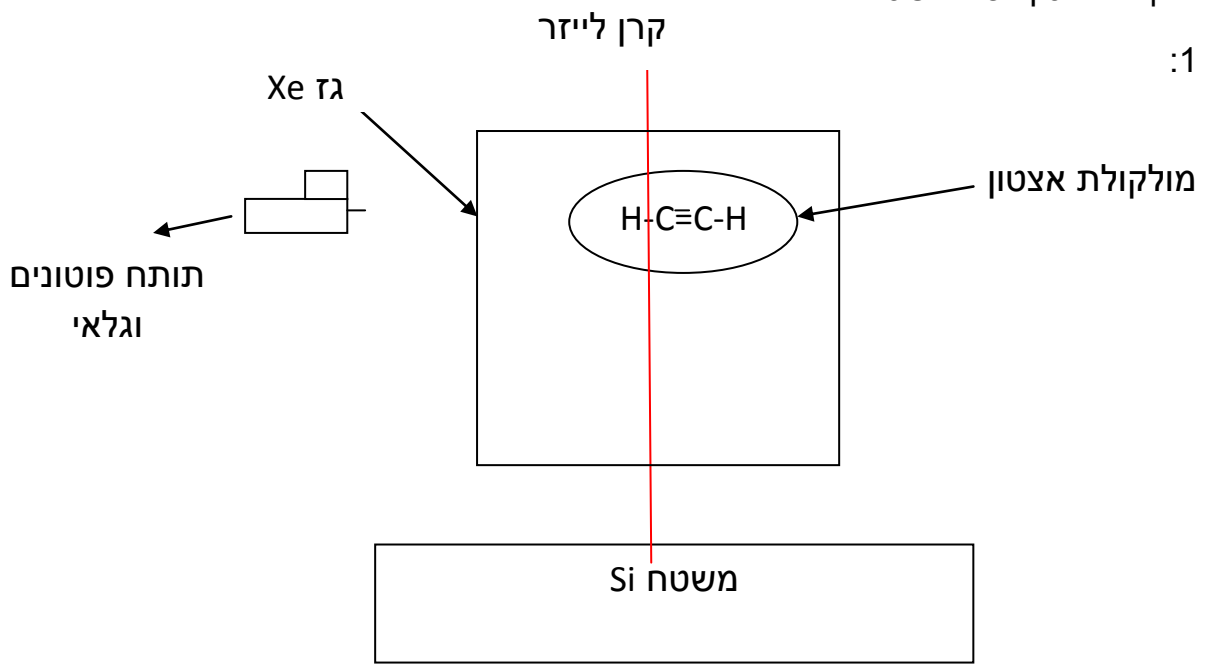
2.א. רקע לפרויקט:

הפרויקט הוזמן על ידי צוות חוקרים מהאוניברסיטה העברית אשר ביקשו לדמות מערכת כימית על ידי סימולציה ממוחשבת בכדי לבחון את השערות הניסוי שבכוונתם לבצע בעתיד. הסימולציה תאפשר לחוקרים לנתח ולהשוות את התוצאות המתקבלות על ידי הסימולציה עם התוצאות המשוערות ובכך תהווה תנאי מקדים לביצוע הניסוי בפועל

2.ב. קצת על הניסוי:

תיאור המערכת הכימית – במערכת יהיה משטח סיליקון (Si) ומעליו הגז האציל קסנון (Xe) במצב צבירה מוצק. במסגרת הניסוי יחממו את משטח הסיליקון בעזרת קרן לייזר. לאחר החימום, משטח הקסנון יתנתק מהסיליקון ויגיע לנקודה בא ישנו תותח פוטונים וגלאי לקריאת הנתונים על האצטון. בנקודה זו תוחדר לאזור הקסנון מולקולת אצטון וישוגרו פוטונים לאזור זה.

איור 1:



התוצאות המשוערות – ההשערה היא ששיגור הפוטונים לאצטון והקסנון יגרום ל H–ים (מימן) במולקולת האצטון להתנתק מה C-ים (פחמן) משום שהם מחוברים בקשר חלש יחסית לקשר בין פחמן – פחמן, ביניהם יכנסו אטומים של Xe ובכך תיווצר מולקולה חדשה H-Xe-C=C-Xe-H.

2.ג. בסימולציה הממוחשבת:

הסימולציה נכתבת בשני שלבים עיקריים:

שלב א':

- סימולציה המדמה מערכת המכילה את משטח הסיליקון, גז הקסנון וחימום המשטח בלבד.
- הסימולציה תהיה ברמת האטומים לדיוק גבוה בתוצאות הניסוי.

חלק זה של הניסוי כבר בוצע בעבר על ידי החוקרים מהאוניברסיטה העברית כאשר ההתבוננות במשטח הסיליקון הייתה ברמת החומר כיחידה אחת וחימומו חושב בעזרת פונקצית החימום של המשטח. חישוב הכוחות הפועלים על האטומים, מיקומם ומהירותם תיתן דיוק גבוה יותר מהגרסה הקודמת.

- בשלב הזה יחושבו הכוחות הפועלים בין זוג סיליקונים, שלשת סיליקונים, זוג קסנונים והכוחות בין סיליקון לקסנון וקסנון לסיליקון (Si-Si, Si-Si-Si, Xe-Xe, Xe-Si, Si-Xe).

חישוב הכוחות נועד לחישוב מיקומם ומהירותם של האטומים בכל שלב. על ידי ידיעת מיקומם של האטומים ניתן לראות מה קורה למערכת הכימית בכל זמן נתון.

שלב ב':

- בסיום הפרויקט הסימולציה צפויה לדמות את המערכת הכימית במלואה כפי שתוארה בעמוד הקודם (איור 1) ובנוסף למערכת משלב א' תיווסף מולקולת האצטון.
- ההתבוננות באצטון תהיה גם כן ברמת האטומים, כלומר, מימן בנפרד מפחמן.

התבוננות זו תאפשר לגלות אם המולקולה שלמה או שהמימן התנתק מהפחמן על ידי חישוב המרחקים ביניהם וכמו כן לדעת האם אטום קסנון נכנס ביניהם ובכך לאשש או להפריך את השערות הניסוי.

סטודנטים נוספים העובדים על הפרויקט:

- רז כהן – עובד על האלגוריתמים לחלק ב' של הפרויקט.
- צללפונית יצחק - עובדת על פונקציות החימום בפרויקט.

3. הבעיות והצעות לפתרון

נתבונן בשתי בעיות עיקריות:

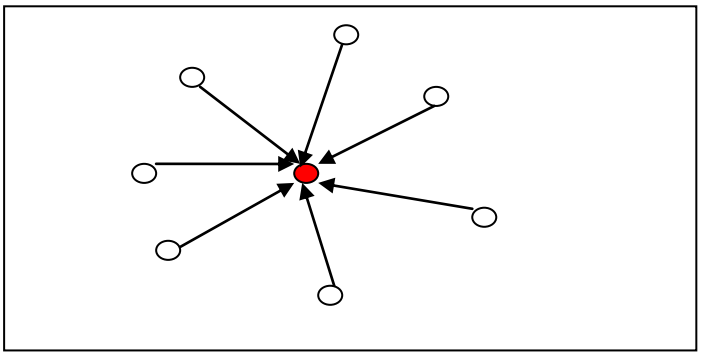
3.א. בעיית זמני הריצה:

סיבה מקדימה לבעיה: ההתבוננות במערכת תהיה ברמת האטומים.

הבעיה: עקב ריבוי האטומים במערכת, חישוב הכוחות על כל האטומים לוקח הרבה מאוד זמן ריצה.

הסבר לבעיה: נתבונן בדוגמא של סך הכוחות הפועלים על אטום קסנון בודד במערכת.

איור 2:



איור 2 מדמה חישוב כוח שכל אטום לבן מפעיל על האטום האדום. סכום כוחות אלה הוא הכוח הפועל על האטום האדום.

מכאן ניתן להסיק שזמן ריצת חישוב כוח על כל אטום הוא O(n) ומכיוון שהחישוב מבוצע על n אטומים, זמן הריצה יהיה O(n^2).

כאשר מדובר באטום סיליקון, בנוסף לחישוב הזוגות מתווסף גם חישוב בין שלשות של סיליקונים. כלומר על אטום סיליקון i נחשב את הכוח שאטומי סיליקון j ו- k מפעילים עליו. במקרה זה זמן הריצה יהיה O(n^3)כאשר n הוא מספר הסיליקונים במערכת.

פתרון לבעיה: משום שמדובר במערכת פיזיקלית, ישנו מרחק מקסימאלי בין שני אטומים i ו – j כאשר במרחק גדול יותר ממנו הפוטנציאל בין i ל- j שואף לאפס בקירוב טוב ולכן הכוח ש- i מפעיל על j גם הוא שואף לאפס ולחלופין, הכוח ש j מפעיל על i.

מכאן כי כל האטומים שרחוקים מאטום i יותר מהמרחק המקסימאלי כלל לא רלוונטיים לחישוב הכוחות שפועלים עליו. לשם צמצום זמן הריצה באופן משמעותי נחלק את מערכת הצירים לתאים בגודל קבוע. כעת נניח כי אטום i נמצא בתא (2,2,2) ונסתכל אך ורק על האטומים שנמצאים בתאים השכנים של התא שלו (כולל התא שהוא נמצא בו). כלומר (1,2,2),(2,2,2),(3,2,2), (1,1,2),(2,1,2),(3,1,2),
(1,3,2),(2,3,2),(3,3,2).... סה"כ 27=3^3 תאים.

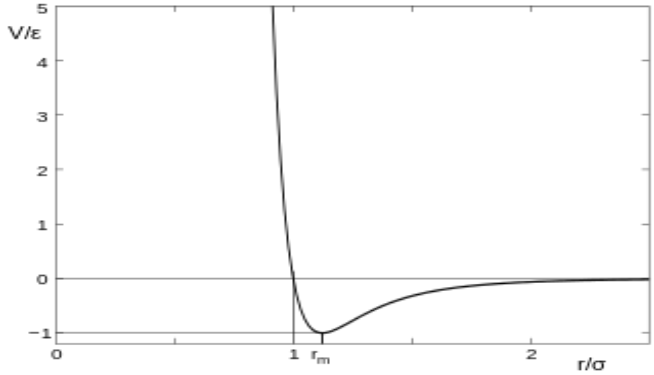
כעת יחושבו רק הכוחות שמפעילים האטומים ה"שכנים" ולא כל האטומים במערכת. בזכות העובדה כי צפיפות האטומים נמוכה, מספר האטומים ה"שכנים" לכל אטום קטנה בהרבה מכמות האטומים. נסמן את המספר המקסימאלי של ה"שכנים" ב MAX_NEIGHBORS ונראה כי זמן הריצה לחישוב הכוחות יורד מ – O(n^2) ל – O(n*MAX_NEIGHBORS) (מאוד קרוב ל - O(n)), ומ – O(n^3) בחישוב שלושה סיליקונים (כאן n הוא מספר הסיליקונים) ל – O(n*(MAX_NEIGHBORS^2)) וגם כאן זה מאוד קרוב ל – O(n).

פתרון זה ניתן ליישום בזכות מודלים מתמטיים פשוטים המתארים את האינטראקציה בין אטומים בעזרתם ניתן לחשב בקירוב מאוד טוב את הפוטנציאלים והכוחות הפועלים על האטומים.

פוטנציאלים הממומשים בסימולציה זו הם:

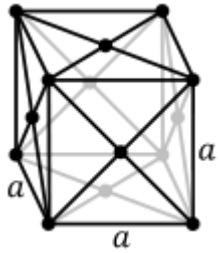
1. פוטנציאל לנארד-ג'ונס לחישוב פוטנציאל בין שני אטומי קסנון והפוטנציאל בין אטום קסנון וסיליקון.
2. פוטנציאל מורס לחישוב פוטנציאל בין שני אטומי סיליקון.
3. פוטנציאל בין שלושה אטומי סיליקון.

נתבונן בפוטנציאל לנארד-ג'ונס:

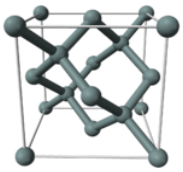


$$V(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

כאשר אפסילון קובע את עומק בור הפוטנציאל כפי שנראה בגרף. כאשר המרחק r בין שני האטומים שווה לסיגמה, הפוטנציאל מתאפס ותחתית בור הפוטנציאל מתקבלת כאשר r = 2^(1/6)*sigma. האיבר (1/r)^12 מתאר את הדחייה ו (1/r)^6 מתאר משיכה. בזכות הצורה הפשוטה של קוביית הקסנון



כאשר a = 2^(1/6)*sigma_xenon*(sqrt(2)) והעובדה כי הפונקציה שואפת לאפס כבר במרחק קטן יחסית נוכל לחשב בקלות מהו המרחק המקסימלי שממנו והלאה נרצה להתייחס לפוטנציאל ולכוח בתור אפס.

בדומה לכך בוצעו החישובים עבור קוביית הסיליקון  ופונקציות הפוטנציאל של הסיליקון.

מלבד לכך שפונקציות אלה פשוטות למימוש, הן גם קצרות וכמו כן גם הנגזרות שלהן קצרות ופשוטות למימוש. מינוס הגרדיאנט של הפוטנציאל שווה לפונקציית הכוח ואלה כמובן הפונקציות שמחושבות בכל איטרציה של התוכנה, לכן הפשטות שלהן מקטינה את זמני הריצה לעומת מודלים אחרים בהם משתמשים בכימיה ולא בסימולציות ממוחשבות.

ייעול נוסף של הסימולציה יהיה להסב את כל חישובי הכוחות בסימולציה לכרטיס מסך מה שאמור לזרז את החישובים פי כמה בזכות ארכיטקטורת התכנות המקבילי לכרטיסי המסך של NVIDIA. הסבה זו תבוצע בשלב ב' של הסימולציה ולכן אין עדיין נתונים על שיפור זה. בעמוד 16 מצורף נספח שמדבר קצת על הארכיטקטורה.

3.ב. בעיית שגיאות החישוב:

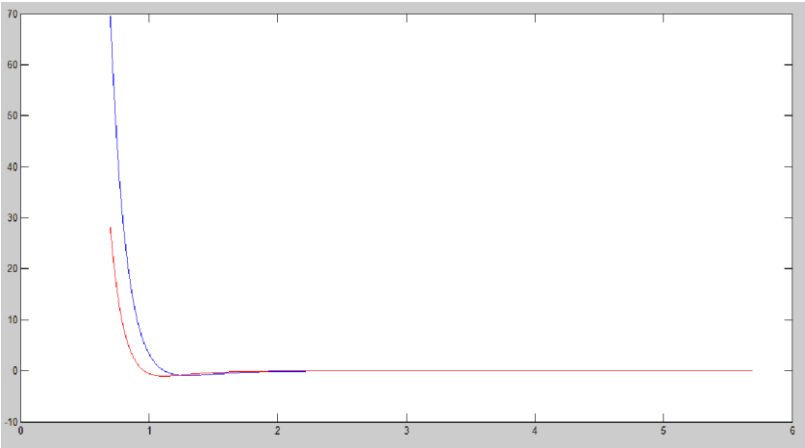
עקב קיזוזי מספרים במחשב ייתכנו שגיאות בתוצאות המתקבלות בכל איטרציה בסימולציה. כמובן שככל שהסימולציה נדרשת לרוץ יותר זמן, השגיאה גדלה, דבר שעלול להשפיע על תוצאות הניסוי ולהביא לתשובות שגויות.

פתרון מניעתי:

על מנת להקטין את הסיכויים לשגיאה, אני שומר על העקביות של סוגי הפרמטרים וביצוע פעולות אריתמטיות על סוגים זהים של פרמטרים. לדוגמא, כתיבת double x = 1/3 במקום double x = 1.0/3.0 תגרור שגיאה של 0.333333333... כי 1 ו 3 מתייחסים ל int והתוצאה שתתקבל תהיה 0 במקום שליש.

פתרונות:

פתרון ראשון טמון בפונקציות הפוטנציאלים. לאחר כתיבת פונקציות הפוטנציאלים והכוחות, כל הפונקציות נבדקו והודפסו גרפים של הפוטנציאל (חלקי אפסילון שהוא קבוע נתון) כפונקציה של המרחק (חלקי סיגמה שהיא גם כן קבוע שנתון).



הבדיקה נעשתה במטרה לבדוק את נכונות השיטות אל מול הנתונים בספרות ונקבעה שגיאה מותרת של עשר במינוס 17. בדיקת הפוטנציאל והכוח של שני אטומי סיליקון העלתה כי ישנה שגיאה בגודל עשר במינוס 10. כלומר היכן שהפוטנציאל חלקי אפסילון היה אמור להיות 1- התקבלה תשובה 0.9999999998- וכמו כן הכוח בנקודה זו לא התאפס ממש. לכן הוחלט להשתמש בפונקצית פוטנציאל בעלת קירוב טוב יותר שניתנה על ידי פרופ' תמר עוז, פוטנציאל מורס שנמצא מדויק הרבה יותר.

פתרון נוסף שמומש הוא קירוב המיקומים החדשים על ידי טורי טיילור. נמצא כי הנוסחה למציאת מיקום חדש x=x0+v0*t+(a*t^2)/2 לא הייתה מדויקת מספיק ולכן פותחה נוסחה עם נגזרת שנייה ושלישית לפי t.

4. תיאור אב טפוס:

4. א. תיאור אב-טיפס ומימוש הפתרונות:

הסימולציה נכתבת בשפת ++C על מערכת הפעלה Windows 7 בסביבת עבודה של

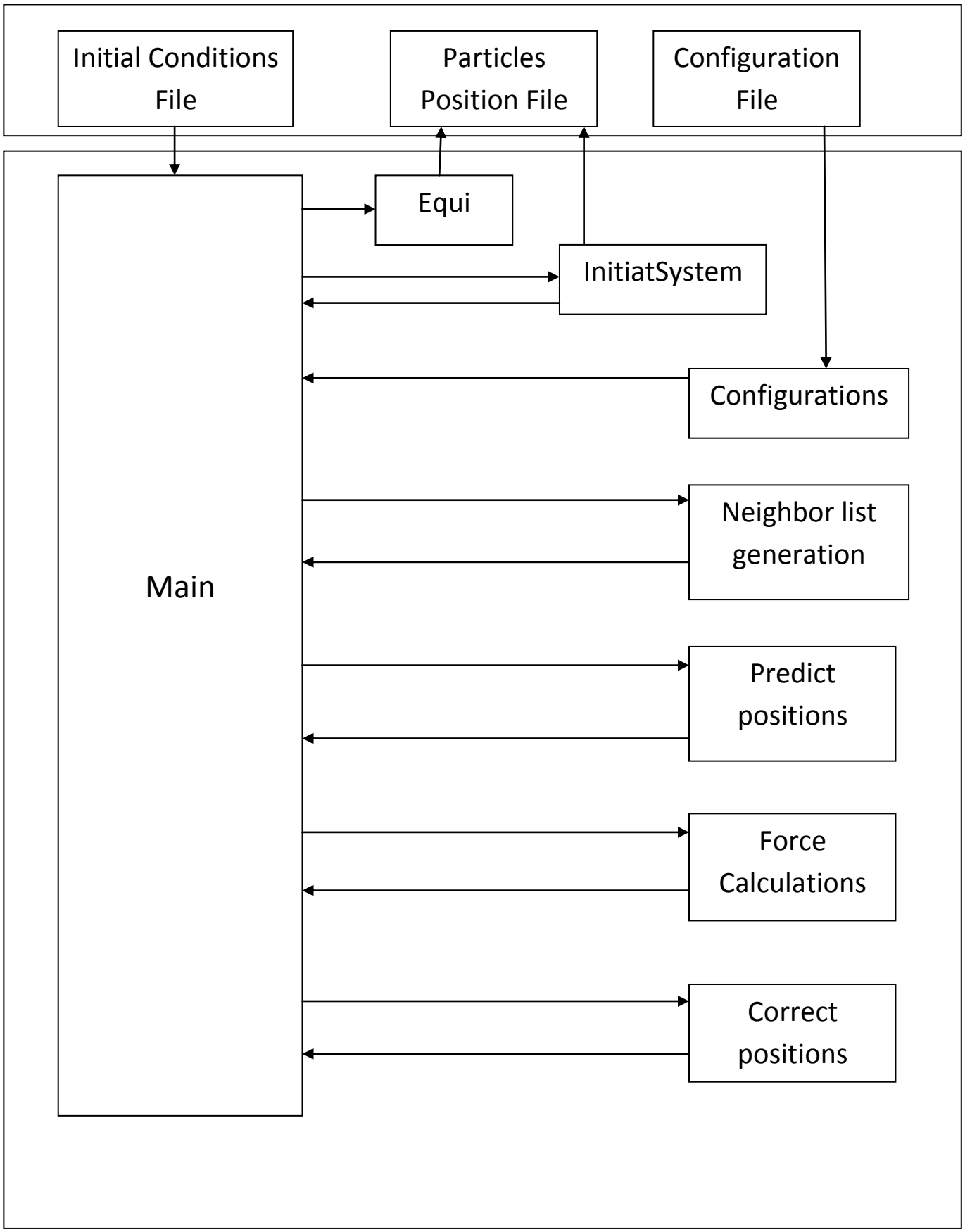
Microsoft Visual Studio Ultimate 2012.

הסימולציה מחולקת לארבעה חלקים עיקריים:

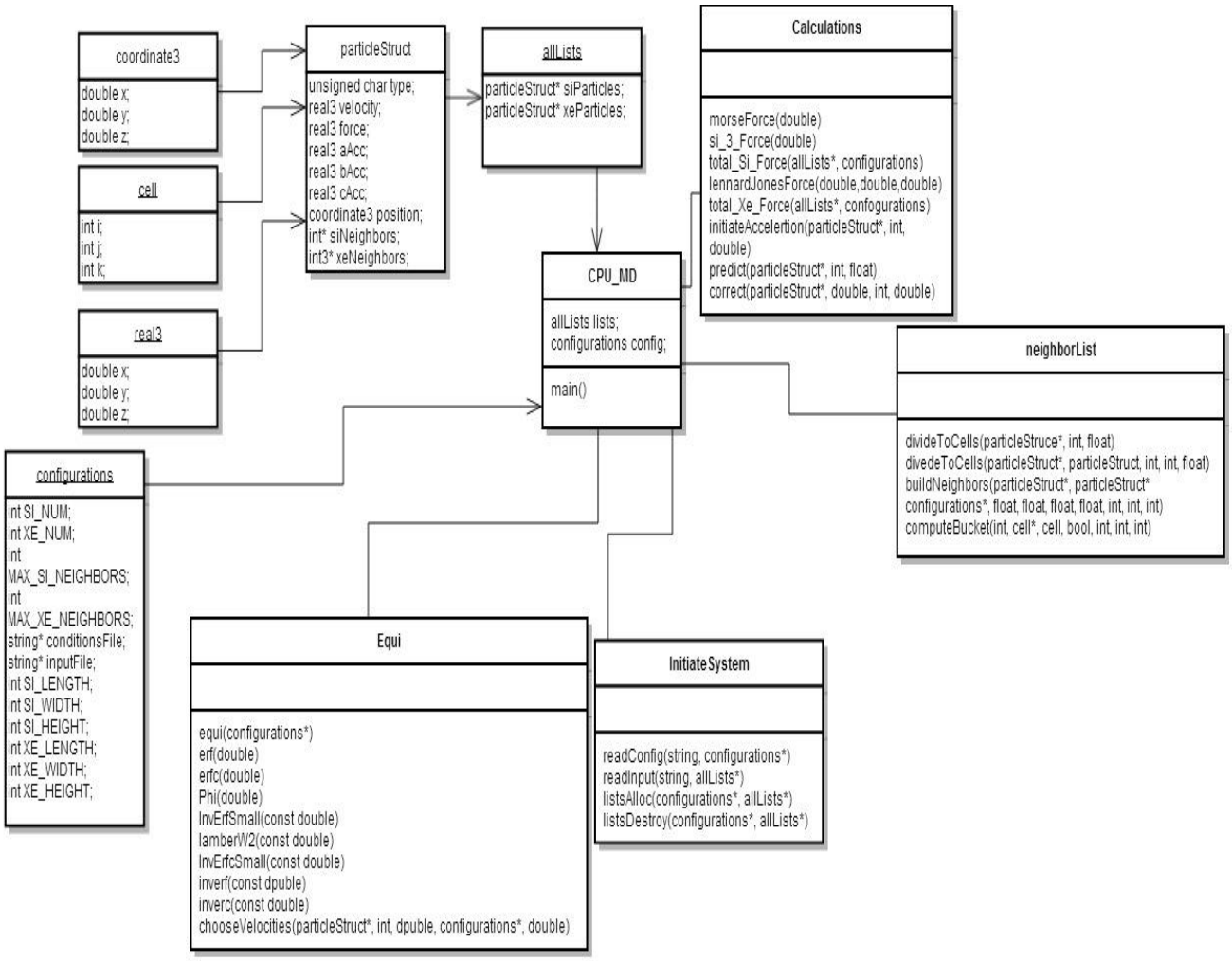
1. חישובים – מחלקה אחת calculations המכילה את כל השיטות שמחשבות את הפוטנציאלים, כוחות ומיקומים, מהירויות ותאוצה חדשים.
2. בניית רשימות שכנים – מחלקה אחת האחראית לחלק את כל המערכת לתאים המתאימים ולשייך כל אטום לתא שלו (לפי המרחקים המתוארים בפתרון בעיית זמני הריצה - לכל חומר במערכת תאים הגדלים שונים). לאחר החלוקה, נקראת השיטה שבונה לכל אטום את רשימות השכנים שלו. לכל חומר יש מערך מבנים משלו, כל מבנה מייצג אטום ויש לו - מיקום, מהירות, תאוצה, התא אליו הוא שייך, רשימת שכני סיליקון ורשימת שכני קסנון. לכן נקרא לשיטה בשביל לבנות שכני סיליקון לכל אטום בשתי הרשימות ושכני קסנון לכל אטום בשתי הרשימות.
3. אתחול המערכת – שתי מחלקות עיקריות Equi, InitiateSystem האחראיות על איתחול המערכת הכימית, סידור האטומים במיקומים התחלתיים נכונים, הקצאת זיכרון ואתחול המערכים של האטומים, מתן מהירויות התחלתיות ולבסוף שחרור הזיכרון.
4. מחלקה CPU_MD המכילה את ה main ואחראית על קריאת השיטות המתאימות בכל זמן ויצירת חלון הסימולציה שנעשית עם OpenGL.

תרשים מודולים אלה בעמוד הבא.

4.ב.



ג.4. תרשים מחלקות וסטרוקטורות עיקריים:



5. בדיקות:

בשלב זה נבדקו פונקציות הפוטנציאלים והכוחות, רשימות השכנים, המהירויות ההתחלתיות וכמו כן חוק שימור האנרגיה במערכת וזמני ריצה בלי רשימות שכנים לעומת זמני ריצה עם רשימות שכנים.

נכונות תוצאות:

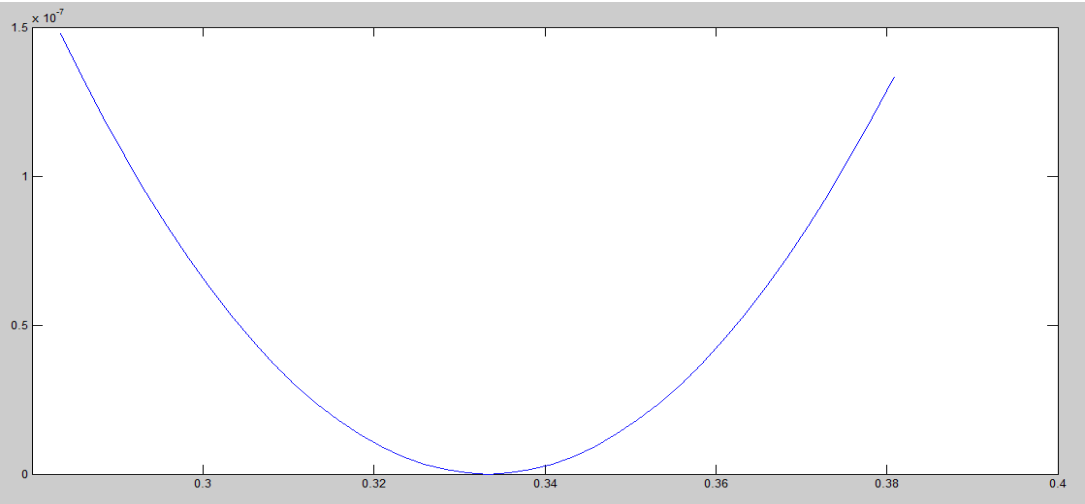
הגרפים שהתקבלו עבור הפוטנציאלים והכוחות הדו גופיים הם הגרפים הצפויים לפי הספרות וכפי שתוארו קודם.

הגרפים שהתקבלו עבור פונקצית הפוטנציאל והכוח של שלושה אטומי סיליקון נכונים גם כן.

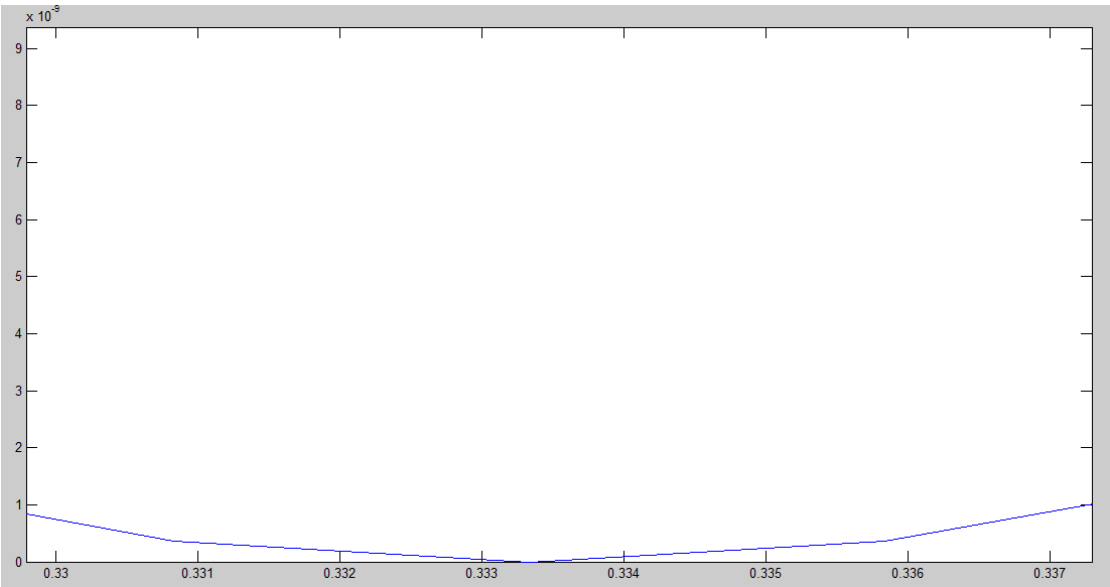
נוסחת הפוטנציאל של שלושה אטומי סיליקון מתאפסת כאשר (cos(x) + 1/3) מתאפס, כלומר x = arcos(1/3) = 109.4712206. הקוד לבדיקה:

```
coordinate3 m;
coordinate3 i;
coordinate3 j;
m.x = 0.0;
m.y = 0.0;
m.z = 0.0;
j.x = pow(2.0,(1.0/6.0))*sigma_Si;
j.y = 0.0;
j.z = 0.0;
i.z = 0.0;
double cosal;
for(int k = -20; k < 20; k++)
{
        cosal = (1.0/3.0)+((double)k/400.0);
        i.x = -cosal*pow(2.0,(1.0/6.0))*sigma_Si;
        i.y = sqrt(1.0-(cosal*cosal))*pow(2.0,(1.0/6.0))*sigma_Si;
        file<<((v3_derivative_of_rix(m, j, i, distance2(m,j),
        distance2(m,i),distance2(i,j)))+(v3_derivative_of_rix(m, i, j,
        distance2(m,i), distance2(m,j),distance2(i,j))))/epsilon_Si;
}
```

הגרף הבא מתאר את הפוטנציאל חלקי אפסילון כפונקציה של מינוס קוסינוס הזווית:

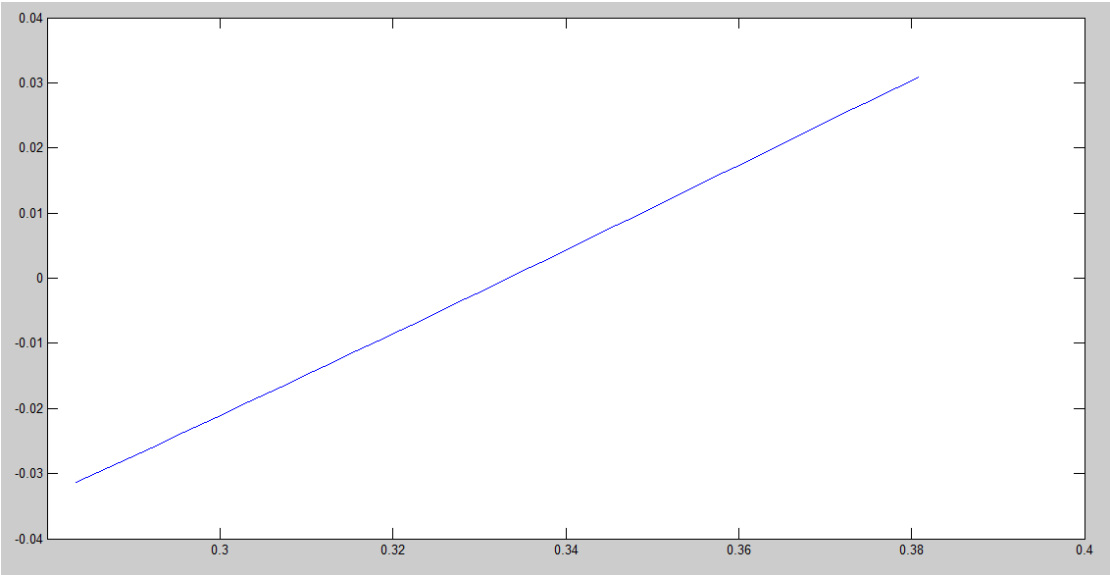


ובהגדלה:

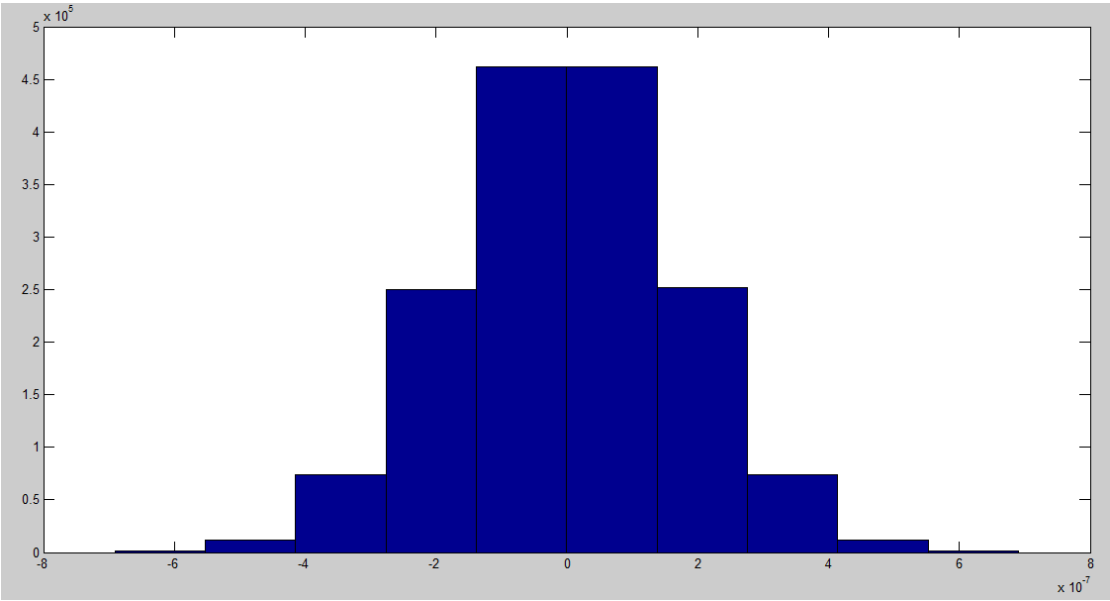


ניתן לראות שנכונות הפונקציה אוששה והפוטנציאל התאפס בדיוק בשליש (בגלל שציר ה X הוא מינוס קוסינוס אז אפשר להבין שהפוטנציאל מתאפס במינוס שליש).

נגזרת הפוטנציאל גם היא התקבלה בדיוק כפי שמצופה מבחינה מתמטית:



היסטוגרמת המהירויות שהתקבלה בכל ציר גם מתפלגת בכל ציר סביב האפס בדיוק כפי שנדרש מהמערכת לתוצאות נכונות.



זמני ריצה:

זמני הריצה נבדקו עבור 2000 אטומי קסנון (5 שכבות לאורך ולרוחב ו – 10 שכבות לגובה) ונלקח זמן ממוצע עבור 10 הרצות של 200 איטרציות.

תוצאות:

עם רשימות 16~ שניות.

בלי רשימות שכנים – 656~ שניות (קרוב ל 11 דקות!!).

נמצא כי עם רשימות השכנים זמן הריצה התקצר פי 41.182.

6. פירוט טכני

- הסימולציה לגרסת המעבד תכתב תוך כדי התחשבות בכך שהיא חייבת להיות גמישה למעבר לגרסת כרטיס המסך. הסיבה לכך נובעת מכך שהרבה חישובים יכולים להתבצע במהירות הרבה יותר גדולה על גבי כרטיס גרפי בהשוואה למעבד רגיל ולכן הפרויקט מכוון לגרסת הכרטיס הגרפי ושפת התכנות וסביבת העבודה נבחרו עקב סיבה זו.

שפת תכנות, כלי תוכנה וסביבת עבודה:

- הפרויקט נכתב בשפת C++ ב - Visual Studio 2012.
- מערכת הפעלה – Microsoft Windows 7 64-bit.
- גרסת ה - CPU נבדק עם מעבד i3-2328M.
- גרסת ה - GPU תיבדק עם NVIDIA GeForce GTX 560 Ti בעל 384 ליבות CUDA.
- מהדר ל CUDA – nvcc (NVIDIA Cuda Compiler).
- גרסת ה GPU תכתב עם שימוש ב CUDA Runtime API (יותר ידידותי לתכנות מ CUDA Driver API ואין ביניהם הבדלי ביצועים מוכחים).

7. סיכונים לשלב ב':

סיכון	הסתברות שיקרה	השפעה	צעדים למניעה	תוכנית ב'
שימוש בסטורקטורה לא נכונה לביצוע החישובים עבור מולקולת האציטון	בינונית	גבוהה	בדיקת נכונות חישובי הכוחות והמרחקים של אטומי הקסנון באינטראקציה עם אטומי האציטון עם סטרוקטורות שונות לפני בניית הסטרוקטורה	מעבר לסטרוקטורה יותר מתאימה
חישובים שגויים עקב קיזוזי מספרים במחשב – סיכון עיקרי (חישובים לא נכונים לא יענו על דרישות הפרויקט)	גבוה	גבוהה	מימוש האלגוריתמים לפי חוקי התכנות המדעי הקיימים	אין
מימוש לא יעיל למעבר לגרסת ה – GPU	בינוני	בינונית	הפרויקט נכתב במטרה למעבר לגרסת כרטיס המסך ולכן נלקחים בחשבון הכללים המנחים לתכנות מסוג זה גם אם זה עולה ביעלות פחותה בגרסת המעבד	שינוי תיכון הפרויקט להתאמה טובה יותר לגרסת ה – GPU

8. לוח זמנים שלב ב':

משימה	תאריך משוער לסיום
התממשקות לפונקצית חימום הלייזר	15/03/2013
מימוש פונקציות לחישוב כוחות	22/04/2013
התממשקות לפונקצית חימום הפוטונים	29/05/2013
חישוב מיקומים ומהירויות חדשים של האטומים	05/06/2014
שילוב כל השלבים, בדיקות אחרונות	12/06/2014
מעבר לגרסת GPU וכתיבת דו"ח מסכם סמסטר ב'	01/07/2014

ספרות מקוצעית:

VNU Journal of Science, Mathematics – Physics 24 (2008) 125 -131 - http://tapchi.vnu.edu.vn/tl_3_08/Hung.pdf

PHYSICAL REVIEW B 75, 155207 (2007) - http://www.virginia.edu/ms/research/wadley/Documents/Publications/Bond_Order_Potential_For_Silicon.pdf

http://courses.physics.illinois.edu/phys466/sp2013/projects/2000/team8/node3.html

נספחים:

קצת על תכנות מקבילי על גבי GPU

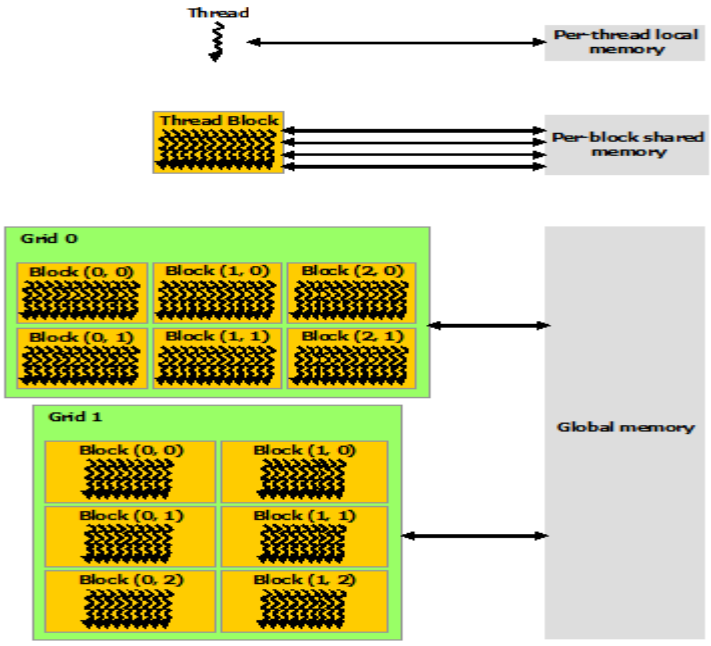
נקרא גם תכנות CUDA.

CUDA - היא ארכיטקטורת מחשוב מקבילי שפותחה על ידי חברת NVIDIA כדי לאפשר לפתח ולהריץ תוכנות בעזרת GPU.

אלמנטים בתוכנה:

- Kernels - הפונקציות שמבצעות את החישובים העיקריים בתוכנה שצפויים לקחת הרבה זמן ריצה.
- קוד המעטפת שאחראי על ניהול זיכרון, העתקת יחידות מידע מהמחשב המארח לכרטיס הגרפי ולהיפך ואחראי לקריאת ה Kernels.

התרשים הבא מתאר את הזיכרון בהתקן:



נשים לב ל – Block ול - Thread . בקריאה ל Kernel מציינים את כמות הבלוקים וה - Thread-ים בהם אנו מעוניינים להשתמש ובעצם כל Thread מהווה תהליכון. ה – Kernels בפרויקט זה יעבדו עם מערכים גדולים כאשר על כל אינדקס במערך יעבוד Thread נפרד ובכך ייווצר מקבול התוכנה וזמני הריצה יצומצמו משמעותית.