**המחלקה להנדסת תוכנה**

**פרויקט גמר – תשע"ט**

**סימולציות על אטומים**

### **Molecular Dynamics Simulator**

**מאת**

**אופק ברזני**

**מנחה אקדמי: דר' יהודה חסין אישור: תאריך:**

**רכז הפרויקטים: דר' אסף שפיינר אישור: תאריך:**

מערכות ניהול הפרויקט:

|  |  |  |
| --- | --- | --- |
| # | מערכת | מיקום |
| 1 | מאגר קוד | github.com/ofekba/project |
| 2 | יומן | https://github.com/ofekba/project/wiki/Meetings-diary |
| 3 | סרטון גרסת אלפא |  |

# מבוא

סימולציה הינה חיקוי של מציאות מורכבת באמצעות מודל מתאים ומטרתה לייצג מאפיינים מסוימים בהתנהגותה של מערכת.

בכימיה החישובית, סימולציה של דינאמיקה מולקולרית (MD) נועדה לדמות מודל תנועה והתנהגות של אטומים בודדים במולקולה או מולקולות לפי בחירה ע"י החלת חוקים של מכניקה קלאסית למציאת פרמטרים כגון מהירות, מטען חשמלי, מיקום, וכו'... פר אטום. השימוש במכניקה הקלאסית לתיאור האטומים מסתמך על פרמטרי forcefields שחושבו או נמדדו מראש (פרמטרים ידועים מראש/ מדידות מניסוי קודם/ לקוחים מהספרות).

מתא-דינמיקה (Metadynamics ) הינה שיטת סימולציה ממוחשבת המשמשת להערכת האנרגיה החופשית ופונקציות אחרות של המערכת בסימולציות מסוג דינמיקה מולקולרית.

הרצת סימולציה מסוג מתא-דינמיקה של דינאמיקה מולקולרית מאפשרת חיקוי תנאי מעבדה אמתיים ומטרתה לייצג מאפיינים בהתנהגותה של מערכת אטומית בעת הפעלת מניפולציות כימיות על המערכת כדי לקבל תחזית על התנהגות המערכת בהשפעת אותן מניפולציות.

LAMMPS הינה תוכנה מבוססת קוד פתוח להרצת סימולציות מתא-דינמיות של MD (דינמיקה מולקולרית) על אטומים ומולקולות בזמן ריצה יעיל בעיקר על מחשבים התומכים בחישוב מקבילי. בניית הסימולציות נעשית ע"י הגדרת מספר צעדי זמן כאשר גודל צעד זמן הינו גודל המוגדר מראש. גודל צעד זמן יהיה בערך שניות, כך שהסימולציה מדמה תהליך כולל של שניות בממוצע.

על מנת להריץ סימולציה בLAMMPS יש ליצור קובץ מסוג dat המכיל את המצב ההתחלתי של האטומים-מידע על האטומים המשתתפים בתהליך הכימי (סוגם, מסותיהם ומיקומם כווקטור מדרגה 3) או ע"י קובץ restart בינארי שהוא תוצר של ריצה קודמת, קובץ field המכיל את הפרמטרים של ה- forcefields וקובץ הסקריפט מסוג in (ייחודי ל( LAMMPS המכיל פקודות של התהליכים הכימיים אותם נרצה להפעיל בסימולציה.

# תיאור הבעיה

*אורך סימולציה* ב LAMMPS הינו ברזולוציה של שניות, ולכן לא ניתן להריץ סימולציה של תהליכים וריאקציות כימיות שאורכן גדול מגודל זה, או בכלל באורך של שניות, דקות ואף שעות, ולכן תחום השימוש *ב LAMMPS הינו מצומצם בהתאם.*

*לאחרונה פורסם מאמר פורץ דרך שניסה לתת פתרון ע"י גישה עובדת והראה תוצאות שיכולות לאפשר הרצת סימולציות של פרקי זמן גדולים יותר מ-, כך שLAMMPS יוכל להריץ סימולציות ברזולוציית זמנים גדולה מהרזולוציה הנוכחית, בסדרי גודל שונים, כך שמגוון התהליכים והריאקציות הכימיות שניתן להריץ כסימולציה במערכת יגדל משמעותית.*

*בפרוייקט הזה, אנו רוצים לאפשר הרצת סימולציה של תהליך ריאקציה שאורך כחצי שעה בין המולקולות EPON862, DETDA ולאחר מכן לחקור את תוצאות הריאקציה ומאפייני החומר החדש שנוצר.*

# תיאור הפתרון

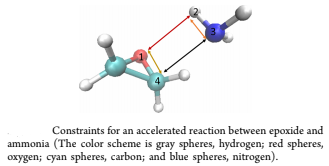
במאמר עליו מתבסס הפרויקט אותו אנסה לשחזר כחלק הראשון של הפרויקט, מתואר אלגוריתם הנועד להוביל לזירוז תהליך צילוב בין 2 מולקולות ליצירת חומר חדש ע"י הפעלת פוטנציאל נוסף על המערכת אם מתרחש אירוע מסוים (יתואר בהמשך). הפעלת הפוטנציאל תוסיף את האנרגיה הדרושה למערכת במטרה להאיץ את התרחשות תהליך הצילוב.

הסימולציה בLAMMPS שנרצה להריץ על מנת לשחזר את תוצאות המאמר מורכבת מארבעה שלבי ריצה במהלכן מופעל על המערכת פוטנציאל בשם REAXFF :

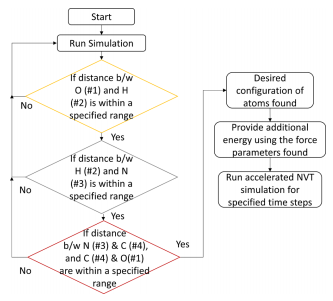
1. MIN- מינימיזציה של האנרגיה.
2. NVT- טמפרטורה ונפח קבועים- בשלב זה נרצה לבצע את הבדיקה להוספת הפוטנציאל הנוסף.
3. NPT – לחץ וטמפ' קבועים.
4. NVE- נפח ואנרגיה קבועים.

הטריגר להפעלת הפוטנציאל הנוסף:

התקרבות אטומים מסוימים בכל מולקולה לאטומים מסוימים אחרים במולקולה השנייה כפי שמתואר בתרשים הבא:



האלגוריתם המפורש שנשחזר על מנת לאתר האם הופעל הטריגר להפעלת הפוטנציאל הנוסף:



נרצה לאתר רביעייה חשודה בין האטומים C, H, O, N הנמצאים במרחקים הבאים:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **C-O** | **N-C** | **H-N** | **O-H** | **PAIR** |
| 1.3-1.6 | 3.0-8.0 | 0.9-1.2 | 1.5-8.0 | **min dist- max dist** (angstrom units) |

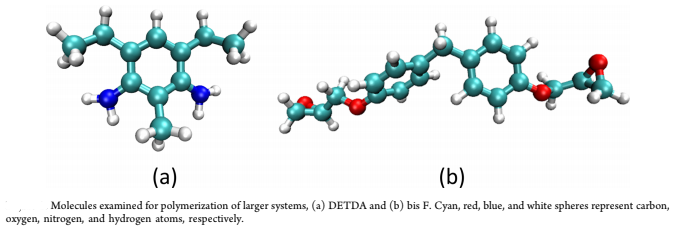
נבצע את הבדיקה הנ"ל בכל צעד זמן. לאחר שאיתרנו רביעייה כנ"ל, נפעיל עליה את הפוטנציאל הנוסף במשך 2000 צעדי זמן. לאחר 2000 צעדי זמן רק הREAXFF ימשיך לפעול עד לסיום הסימולציה.

המולקולות עליהן רצה הסימולציה:

את האלגוריתם נרצה לממש ב-3 דרגות שונות של מולקולות, את הדרגה הראשונה ביצענו בחלק הראשון של הפרוייקט בהצלחה תוך הגעה לתוצאות הרצויות ואת הדרגה השנייה והשלישית נבצע בחלק השני של הפרוייקט.

**דרגה ראשונה-** החלק הראשון של הפרויקט- נרצה לממש את האלגוריתם על 2 מולקולות קטנות יותר מהמולקולות המופיעות במאמר, NH3 ו- C2OH4 , לאתר רביעייה אחת של C, H, O, N להפעיל עליה את הפוטנציאל הנוסף ולבדוק כי אכן לאחר הפעלת הפוטנציאל מתרחשת הריאקציה הרצויה.

**דרגה שנייה-** מימוש האלגוריתם על 2 המולקולות המופיעות במאמר (*EPON862, DETDA )* ביחס 1:2 ואיתור מבין 16 אפשרויות לרביעיות חוקיות שתי רביעיות המקיימות את התנאים לעיל ולהפעיל על כל אחת מהן את הפוטנציאל ולוודא כי אכן בוצעה הריאקציה הרצויה



**דרגה שלישית-** שכפול המולקולות מהדרגה השנייה וביצוע תהליך בדיקת התנאים והפעלת הפוטנציאל על מערכת בסדר גודל משמעותית גדול יותר עם מיליוני מולקולות כנ"ל ביחס 1:2, בדיקה כי אכן הריאקציה מתרחשת, וחקירת תכונות החומר שנוצר. כמו כן הקפדה על זמני ריצה, ושימוש בתכונת החישוב המקבילי של תוכנת הlammps לייעול זמן הריצה.

חישוב הפוטנציאל:

נוסחת לחישוב האנרגיה שמוסיף הפוטנציאל

  
חישוב וקטור הכוחות לכל זוג אטומים i ,j





  
הפרמטרים לחישוב

|  |  |  |  |
| --- | --- | --- | --- |
| **C-N** | **O-H** | **O-C** | **pair** |
| 1.5 | 1.1 | 1.95 |  |
| 300 | 250 | 50 |  |
| 0.75 | 0.75 | 0.5 |  |

בדיקת תקינות התוצאות ושיפורן:

כדי לבדוק את תקינות התוצאות עלינו לבדוק את הריאקציה שהתרחשה- אילו קשרים נוצרו ואילו קשרים נשברו בין כל האטומים במערכת ובפרט מעקב צמוד אחר האטומים המגיבים. אנו נרצה כי יווצרו הקשרים בין N-C, H-O וישברו הקשרים בין H-N, C -O ברביעייה אליה אנו מוספים את הפוטנציאל. לשם כך הוספתי קוד שיוצר קובץ טקסט המכיל פירוט של המרחקים בין כל זוג אטומים במערכת בכל מספר צעדי זמן מוגדר מראש וקוד פייתון שעובר על אותו קובץ ויוצר גרף העוקב בין מרחקים בין כל זוג אטומים רלוונטי להתרחשות הריאקציה כתלות בזמן. מהתבוננות בגרף ניתן להבחין בהתרחשות הריאקציה הרצויה (פירוק והיווצרות הקשרים הרצויים). דרך נוספת לבדיקת הקשרים בין האטומים הינה בדיקת הקובץ טקסט bonds.reax שהוא תוצר של ריצת הסימולציה ועוקב אחרי קשרי הbond של כל אטום. כדי לבדוק שאכן בוצעה שמירה על חוק שימור האנרגיה כתבתי קוד היוצר קובץ טקסט added\_energy.reax ומתעד את האנרגיה שהפוטנציאל הוסיף למערכת בכל צעד זמן, וקוד פייתון שעובר על קובץ זה ומציג גרף של תוספת האנרגיה של הפוטנציאל כתלות בזמן. הגרף עזר לנו לחקור אילו פרמטרים 1F לבחור בחישוב הפוטנציאל על מנת להגיע לריאקציה תוך הקפדה על חוק שימור האנרגיה.

מבנה הקוד:

את הקוד לבדיקת התנאים להוספת הפוטנציאל אוסיף לsource code של LAMMPS כמחלקה מסוג FIX בשפת התכנות C++ בשם fix\_reaxc\_checkFourset. לפירוט מבנה המחלקות בLAMMPS ראו נספח ו'.

כדי להפעיל את הקוד הנ"ל יש להוסיף פקודה מתאימה לקובץ IN המכיל את רצף הפקודות של מהלך הסימולציה בתחביר הבא:

**Fix checkFourset all reax/c/checkFourset 1 dist.reax**

את הקוד לחישוב הכוחות והוספת הפוטנציאל אוסיף כפונקציות למחלקה pair\_reaxc כך שהפוטנציאל הנוסף ירוץ במקביל לפוטנציאל של הreaxff שרץ במערכת.

# סקירת עבודות דומות \ בספרות והשוואה \ סקר שוק

* המאמר עליו מתבסס הפרויקט:

Accelerated ReaxFF Simulations for Describing the Reactive CrossLinking of Polymers Aniruddh Vashisth, Chowdhury Ashraf, Weiwei Zhang, Charles E. Bakis, and Adri C. T. van Duin

Publication Date (Web): July 11, 2018

ישנן שיטות שונות שפותחו כדי לבצע סימולציות בקנה מידה אטומי עבור צילוב של פולימרים (הרצת סימולציות ברזולוציית אורך כולל של שניות כמו בLAMMPS למרות שהתהליך אותו מסמלצים אורכו ברזולוציית דקות-שעות), אך רובן לא סימלצו את כל שלבי ומצבי הריאקציה כולה באופן איכותי. ניסויים למדידת תגובות בקישור בין פולימרים נעשים בדר"כ בטווח של דקות-שעות, טווח שלא מאפשר הרצה בסימולציות בקנה מידה אטומי (כגון LAMMPS ) וכיוון שביצוע ניסויים אלו בזמן אמת דורש עלות גבוהה, פותחו שיטות המבוססות ReaxFF reactive force field.

בשיטה המתוארת במאמר, האטומים המגיבים נמצאים במעקב עד שהם מגיעים לתצורה מסוימת המספקת נקודת התחלה טובה להתחלת ריאקציה. כדי "לעודד" אותם מוסיפים כמות אנרגיה גדולה יותר או שווה למינימום אנרגיה הדרוש להם לצורך תגובה, ובכך להתגבר על המכשול המונע את תהליך הצילוב שיוצר את החומר הרצוי- כלומר זירוז תהליך הצילוב בין החומרים ע"י זיהוי מצב לתחילת הריאקציה והוספת אנרגיה כדי לגרום לתהליך הצילוב להתרחש במידי, אך תוך כדי נשים לב כי לא כל פעם שנפעיל את פוטנציאל האנרגיה הנוסף נקבל את התוצאה רצויה.

בכך אנו מאפשרים הדמיה אמתית של תהליך צילוב בין חומרים בטמפרטורות נמוכות באופן המחקה תגובות כימיות מבלי לאפשר תגובות לא רצויות כתוצאה מטמפרטורה גבוהה. במאמר מתוארת הפעלת השיטה הנ"ל בחקירת תהליך הצילוב בין המולקולות bisphenol F ו- DETDA . התוצאה שהתקבלה הינה שיעור צילוב גבוה יחסית של 82% בין שני המולקולות הללו, ולכן המסקנה הנובעת מהמאמר וכתוצאה מתוצאות ניסויים נוספים המתוארים בו שבוצעו באותה שיטה היא כי שיטה זו היוצרת סימולציות מואצות בREAXFF מהווה כלי שימושי לביצוע סימולציות בקנה מידה אטומי על תהליכים פולימרים שקורים בפועל בזמנים גדולים בהרבה (דקות-שעות) .

בהסתמכות על תוצאות המאמר ומסקנותיו, לאחר שחזור המאמר נרצה להכליל את עקרונותיו כך שנוכל להשתמש בשיטה המתוארת בו לבדיקת בעיות ספציפיות נוספות שמעניינות אותנו עם מולקולות שונות ליצירת חומרים שונים, כלומר נרצה לבצע סימולציה באופן דומה על תהליכי צילוב בין מולקולות אחרות ליצירת חומרים אחרים רצויים וחקר החומרים שנוצרו- תכונות מסוימות כגון מסה, עמידות, חוזק, טמפ' פירוק, וכו'...

* מאמר שני העוסק בפיתוח REAXFF

ReaxFF:  A Reactive Force Field for Hydrocarbons

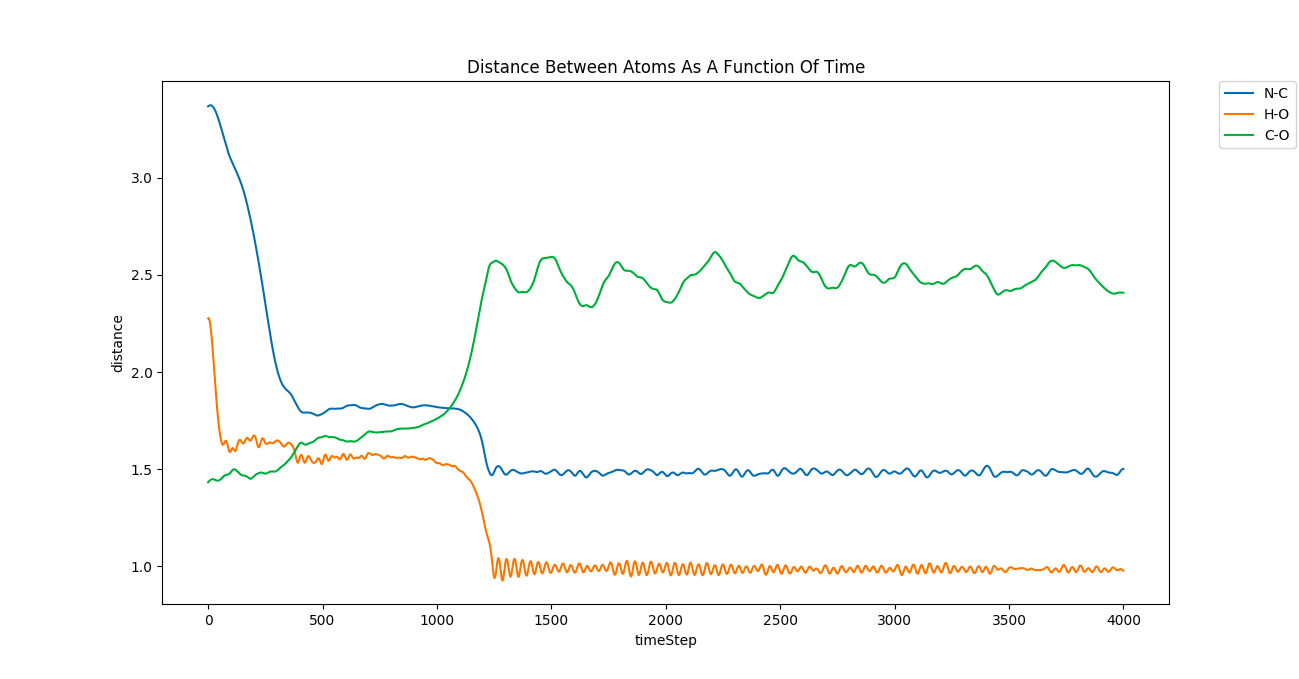
Adri C. T. van Duin‖, Siddharth Dasgupta, Francois Lorant, and William A. Goddard.

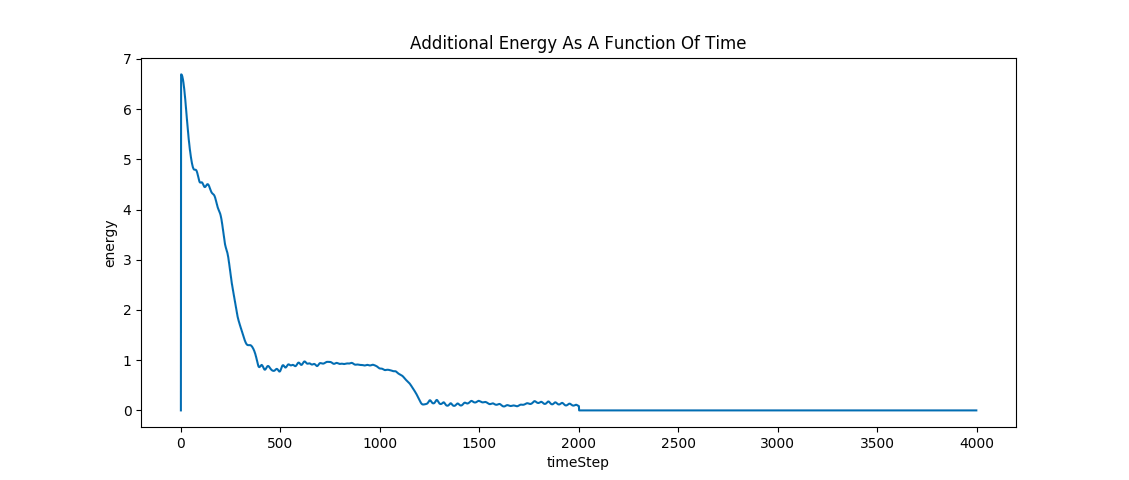
Publication Date: 2011

כדי לאפשר הרצת סימולציית מולקולריות דינאמיות במערכות כימיות תגובתיות בקנה מידה גדול (1000 אטומים או יותר) פותח ReacFF Force Field שדה כוח למערכות ריאקציה כימיות. ReaxFF משתמשים ביחסים הקשורים לתכונות של קשרים כימיים (bonds) בין אטומים– היחס בין bond distance לבין bond order, והיחס בין bond order לבין bond energy והשפעתם על ניתוק הקשר הכימי (הbond) בין אטומים לכדי אטומים נפרדים. כמו כן, ReaxFF מכיל את Coulomb and Morse potentials לשימוש לתיאור אינטרקציות nonbond בין כל האטומים, כלומר אינטרקציות בין אטומים שאינם מחוברים בקשר כימי. הפרמטרים לפוטנציאלים הנ"ל נגזרים מחישובים כימיים קוונטים הנעשים על ניתוקי קשרים כימיים וריאקציות בין מולקולות קטנות.

# סיכום \ מסקנות

בחלק הראשון של הפרוייקט כתבתי קוד המממש את המאמר על מולקולות מהדרגה הראשונה (שתי מולקולות קטנות). כתבתי קוד שמאתר רביעייה חשודה, ומפעיל את הפוטנציאל הנוסף על שלוש זוגות מאותה רביעייה כפי שתואר במאמר ע"י חישוב הכוח שהפוטנציאל מפעיל על כל אטום בכל זוג וכמו כן חישוב וסכימת האנרגיה שהוספת הפוטנציאל גורר לאנרגיה הכללית של המערכת. את הפרמטר 1F בנוסחא לחישוב הפוטנציאל היה צורך לשנות כיוון שהיה מותאם למולקולות הגדולות יותר מהדרגה השנייה, ולכן גרם להתפוצצות המערכת. לאחר חקירת הגרפים שנוצרו עבור כמה ערכי 1F שונים בחרנו ב1.5% מערך 1F המקורי מהמאמר, וסיום הרצת הסימולציה קיבלנו את התוצאות להן ציפינו- אכן התרחשה ריאקציה, הקשרים בין N-C, H-O נוצרו והקשרים בין H-N, C -O התפרקו תוך שמירה על חוק שימור האנרגיה, ושחזור המאמר עבר בהצלחה, משמע אפשר לעבור לשלב הבא בפרוייקט- החלת האלגוריתם על מולקולות מהדרגה השנייה (המולקולות הגדולות יותר המקוריות עליהן נבנה המאמר).





# נספחים

ספרות, תרשימים נוספים, תכנון הפרויקט, טבלת ניהול סיכונים, טבלת דרישות (URD),

## רשימת ספרות \ ביבליוגרפיה

Accelerated ReaxFF Simulations for Describing the Reactive Cross-Linking of Polymers

By: Aniruddh Vashisth, Chowdhury Ashraf, Weiwei Zhang, Charles E. Bakis, and Adri C. T. van Duin

Publication Date: July 11, 2018

Effect of chemical structure on thermo-mechanical properties of epoxy polymers: Comparison of accelerated ReaxFF simulations and experiments

By: Aniruddh Vashisth, Chowdhury Ashraf, Charles E. Bakis, Adri C.T. van Duin

Publication Date: August 11, 2018

ReaxFF: A Reactive Force Field for Hydrocarbons

By: Adri C. T. van Duin, Siddharth Dasgupta, Francois Lorant, and William A. Goddard

Publication Date: March 30, 2001

## תרשימים וטבלאות

**מסכים (אם לא למעלה)**

**תרשימי תיכון כגון: דיאגרמת רכיבים \ הפצה (UML), דיאגרמת ישויות**

**טבלאות במסד נתונים**

# פרוייקט מחקרי יכול להוריד את סעיף ב' -- **באישור המנחה בלבד!**

## תכנון הפרויקט

|  |  |
| --- | --- |
| 8.10. | פגישת היכרות, רקע מקדים על הפרויקט והצגת הבעיה |
| 31.10 | הבנת מבנה התוכנה LAMMPS, מבנה קבצי הקלט לריצות. |
| 26.11 | הבנת מבנה ההיררכי של המחלקות בsource code של LAMMPS, וכמו כן אילו מחלקות דרושות למימוש האלגוריתם |
| 2.12 | כתיבת האלגוריתם שנועד לוודא מצב התחלתי להפעלת הפוטנציאל |
| 30.12 | כתיבת קוד המפעיל את הפוטנציאל על המערכת |
| 20.7 | הרצת הסימולציה ובדיקת התוצאות המתקבלות ותיקון הקוד עד קבלת תוצאות תקינות |
| 27.1 | פגישה עם תמר רז, נעמי מרום והסטודנטים מהנדסת חומרים להצגת תוצאות ותכנון המשך עבודה |
| 22.2 | תחילת עבודה על המולקולות הגדולות |
| אמצע מרץ | הרצת הסימולציה בהצלחה על המולקולות הגדולות |
| תחילת אפריל | תחילת עבודה על שכפול המולקולות הגדולות |
| סוף מאי | חקירת המולקולה החדשה שנוצרה |
| סוף יוני (סוף סמסטר ב') | סיום שיפור זמני ריצה ועריכת הקוד לעבודה במקביליות |

## טבלת סיכונים

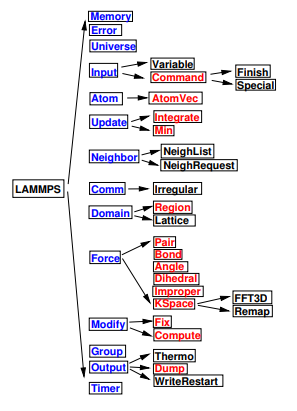
|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **הסיכון** | **חומרה** | **מענה אפשרי** |
| 1 | הערכה שגויה של גודל הפרויקט | גבוה | ארגון מחדש של תכנון העבודה, הפרדה בין חשוב למיותר, סידור עבודה והגדרת מטרות ברורות לכל שלב עבודה |
| 2 | להיכשל בניסיון לשחזור הניסוי במאמר | גבוה מאוד | לנסות להיעזר בנספחים למאמר, לבצע שינויים בקבצי האתחול של האטומים |
| 3 | לא להצליח להכליל את האלגוריתם עבור 2 מולקולות בודדות למספר גדול של מולקולות | גבוה | כתיבה מחדש של האלגוריתם מגישה שונה. |
|  |  |  |  |

## רשימת\טבלת דרישות

**טבלת דרישות (User Requirement Document)**

|  |  |
| --- | --- |
| מס' דרישה | תיאור |
| 1 | זמן הריצה של האלגוריתם יהיה מיטבי |
| 2 | כתיבת האלגוריתם לכרטיס מסך |
| 3 | כתיבת האלגוריתם באופן המאפשר חישוב מקבילי |
| 4 | תוצאות הפעלת האלגוריתם יהיו זהות לתוצאות המתוארות במאמר |
| 5 |  |

## מבנה המחלקות בקוד מקור של LAMMPS



המחלקות בהן אשתמש בכתיבת הקוד-

Neighbor- לצורך שימוש ברשימת השכנויות המתחוזקת במהלך הסימולציה לבדיקת המרחקים בין ארבעת האטומים שנרצה להגדיר כרביעייה חשודה עליה נפעיל את הפוטנציאל הנוסף.

Fix – לצורך הוספת האלגוריתם לזיהוי רביעייה והפעלת הפוטנציאל הנוסף כפקודה שניתן להוסיף לקובץ הקלט in לריצה של סימולציה.

Thermo- הוספת האנרגיה של הפוטנציאל הנוסף לחישוב האנרגייה הכולל של המערכת

Atom- המבנה נתונים המשמש לתיאור האטומים בסימולציה. מכיל בתוכו כמה סוגי מבני נתונים העוקבים אחרי האטומים במערכת, סוג אחד עבור הרצה באופן מקבילי וסוג שני עבור הרצה באופן סדרתי. הקוד שאכתוב צריך לתמוך בשניהם כדי לאפשר הן הרצה באופן סדרתי של הסימולציה על מעבד אחד, והן הרצה באופן מקבילי של הסימולצייה על גבי כמה מעבדים.

כדי להריץ סימולציה באופן מקבילי על כמה מעבדים מחלקים את תיבת הסימולציה לעשרים ושבעה חלקים המחולקים בין המעבדים של המחשב, ולכל חלק רשימת אטומים אותו הוא מכיל (local atoms) ורשימת אטומים אותם תשעת החלקים מסביבו מכילים (ghost atoms) . כאשר נעבוד באופן מקבילי, נרצה לבדוק מרחקים בין אטומים לוקאליים בכל חלק ובכן מרחקים בין הלוקאליים לבין אלו בחלקים מסביב- הghost atoms.

Force- לצורך הוספת חישוב הכוחות של הפוטנציאל לכל אחד מארבעת האטומים עליהם פועל הפוטנציאל, במקביל להפעלת הפוטנציאל reaxff באופן קבוע על המערכת.

הבחירה ב LAMMPS נעשתה מהמניעים הבאים:

בעבר ניסו לפתח מערכת לסימולציות על אטומים במכללה, אך הפיתוח צרך המון זמן ותהליך דיבוג ארוך. לבסוף הוחלט להיצמד לחבילת תוכנה קיימת- LAMMPS ולהרחיב אותה מתוך כוונה להוסיף בעתיד היכולות שונות לחיזוק המערכת. כמו כן LAMMPS הינה מערכת יציבה והעיקרון המנחה אותה- זמני ריצה יעילים מבוסס על חישוב מקבילי מאפשר הרצת סימולציות ביעילות מקסימלית.