

היי נעמי,

ביצעתי הפעם ריצה חדשה של 22K צעדי זמן, במהלכה 10K צעדי זמן ראשונים רק REAX מופעל ללא הפעלת פוטנציאל נוסף ומסתבר שהמרחקים המחוברים בסוגי ריצות שונות (serial or 1/2/4/8/12 threads) עדיין לא אותו הדבר. לאחר מכן מתבצעת איתור רביעייה והפעלת הפוטנציאל הנוסף עליה.

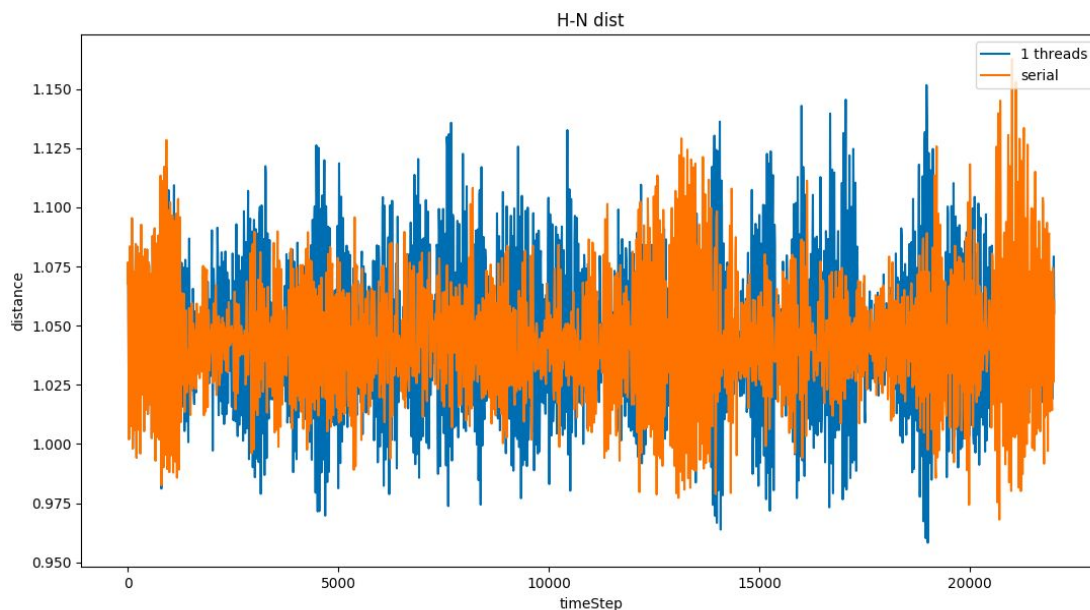
כיוון ש10K צעדי זמן הראשונים לא היו זהים, בכל אחת מהריצות אותה רביעייה אחרת (הבחירה האקראית לא באמת אקראית בגלל שזה מבוסס על ערך SEED מסויים ששתלתי כדי לוודא שאותה בחירה אקראית תתרחש בין סוגי הריצות שונים) וכמו כן

בריצה עם 1THREAD רביעייה חשודה אותה רק בצעד זמן 10.6K. כפי שניתן לראות בתמונה הבאה:

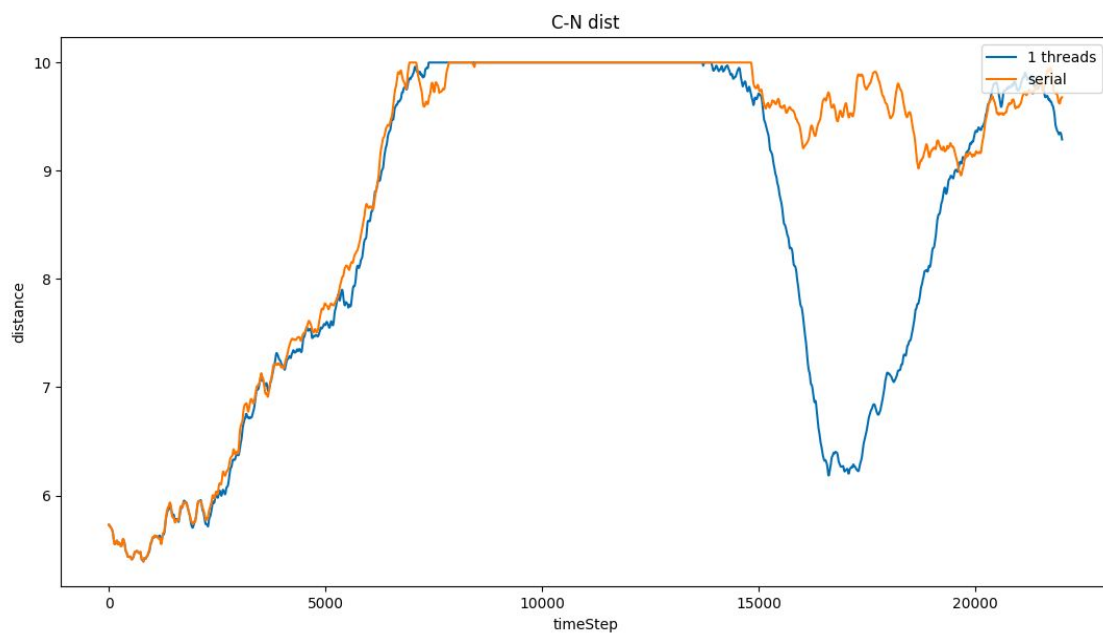
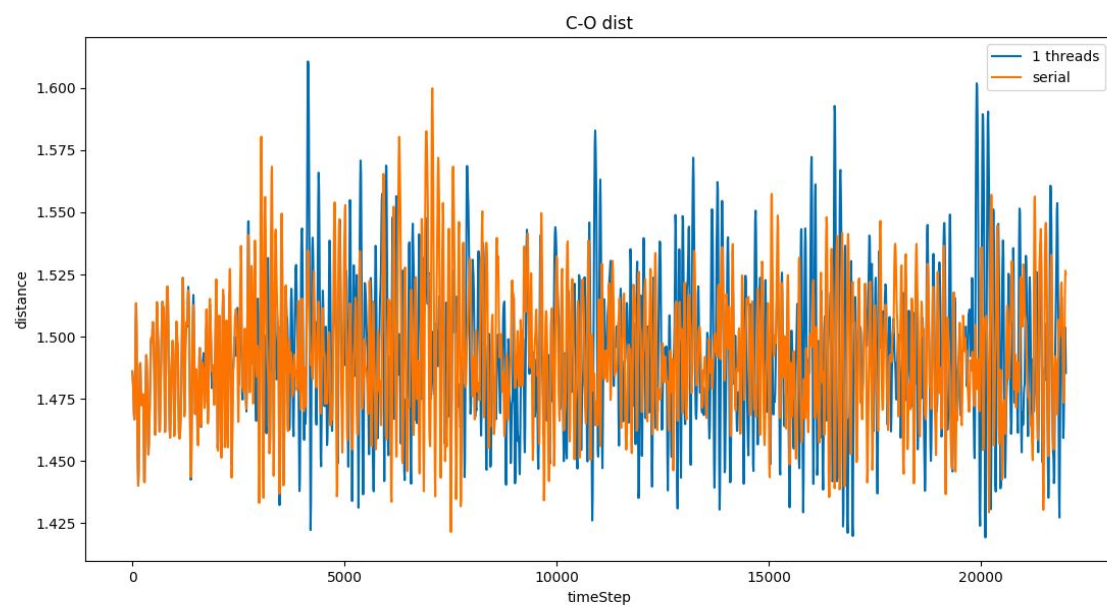
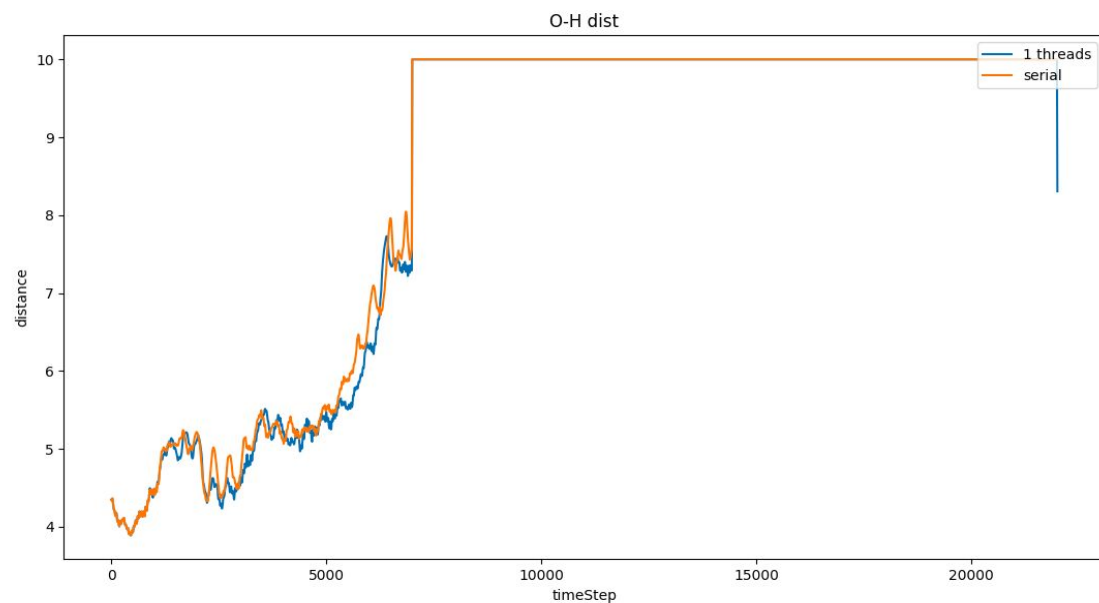
dist_x_.reax הוא שם הקובץ המתעד מרחקים בסוג ריצה עם X THREADS כאשר S מסמל SERIAL, צעד הזמן שבו נמצאה הרביעייה, ושורה מתחת הרביעייה עליה הופעל הפוטנציאל הנוסף.

```
found fourset of dists1.reax at 10000  
['1/1-', '997', '1810', '1788', '1005']  
  
found fourset of dists2.reax at 10000  
['1/1-', '875', '1652', '1632', '879']  
  
found fourset of dists4.reax at 10000  
['1/1-', '1312', '1871', '1850', '1308']  
  
found fourset of dists8.reax at 10000  
['1/1-', '997', '1808', '1787', '1005']  
  
found fourset of dists12.reax at 10000  
['1/1-', '1312', '1840', '1819', '1308']  
  
found fourset of distss.reax at 10630  
['1/1-', '997', '1810', '1788', '1005']
```

ציירתי גרף המשווה מרחקים בין ריצה SERIAL לבין ריצה 1 THREAD כיוון שבשתיהן נמצאה אותה רביעייה בצעד זמן 10K,



ואלו הגרפים שהתקבלו:



ב10K צעדי זמן ראשונים שבהם רק REAX מופעל עדיין המרחקים לא זהים ויש שגיאה כלשהי שגורמת לכך שלא נמצאו אותן רביעיות בכל אחת מסוגי הריצות למרות שהאקראיות הייתה אמורה לבחור תמיד באותה רביעייה אם נמצאה.

<u>The run</u>	<u>serial</u>	<u>1 thread</u>	<u>2 threads</u>	<u>4 threads</u>	<u>8 threads</u>	<u>12 threads</u>
Timestep	10630	10000	10000	10000	10000	10000
The chosen fourset (O H N C)	997 1810 1788 1005	997 1810 1788 1005	66 1407 1385 64	1083 1528 1508 1091	1141 1807 1787 1139	1126 1716 1695 1134
The founded foursets	fourset #0: 997 1810 1788 1005 fourset #1: 997 1809 1788 1005	fourset #0: 997 1810 1788 1005 fourset #1: 997 1809 1788 1005 fourset #2: 997 1810 1788 1005	fourset #0: 66 1407 1385 64 fourset #1: 223 1404 1384 231 fourset #2: 66 1405 1384 64 fourset #3: 66 1404 1384 64 fourset #4: 66 1406 1385 64 fourset #5: 66 1407 1385 64 fourset #6: 1312 1841 1819 1308 fourset #7: 1312 1840 1819 1308	fourset #0: 1083 1528 1508 1091 fourset #1: 1097 1528 1508 1093 fourset #2: 1312 1840 1819 1308 fourset #3: 1312 1841 1819 1308	fourset #0: 1141 1807 1787 1139 fourset #1: 1141 1715 1694 1139 fourset #2: 1141 1808 1787 1139 fourset #3: 1141 1807 1787 1139 fourset #4: 1055 1807 1787 1053	fourset #0: 1126 1716 1695 1134 fourset #1: 1126 1717 1695 1134 fourset #2: 1126 1716 1695 1134 fourset #3: 180 1746 1725 188 fourset #4: 1312 1500 1478 1308 fourset #5: 1312 1841 1819 1308 fourset #6: 1312 1840 1819 1308 fourset #7: 1312 1839 1818 1308 fourset #8: 1312 1499 1478 1308 fourset #9: 1312 1838 1818 1308

כפי שניתן לראות בכל סוג ריצה אותרו רביעיות שונות לגמרי למרות שהקוד שרץ ב10K צעדי זמן ראשונים הוא REAX בלבד ללא שום התערבות מהקוד שלי. ולכן גם הגרפים של המרחקים נראים שונים לגמרי בין סוגי הריצות.

מאת: Naomi Rom
אל: Ofek Barazani
עותק: Yehuda Hassin
תאריך: 10 ביוני 2019, 10:13
נושא: Re: בדיקת תקינות קוד מקבילי

הי אופק,

בריצה החדשה ששלחת אתמול - כפי שכתבת, הרביעיה שאותרה בסריאלי זהה לזו שאותרה עם thread יחיד. אם הבנתי נכון - לשם ההשוואה בין החישובים השתמשת באותו SEED לכל הריצות (ולא השארת את ה-SEED לבחירה אקראית), וזה טוב מאד לבדיקת הקוד וההשוואה שאת עושה.

נראה לי שיתכן ומבחינה נומרית נכנס איזה רעש אקראי לחישובים כתלות במספר ה-threads, דבר שמסביר שכן צפוי שיהיו הבדלים קטנים בין החישובים, שיובילו לבחירה שונה של רביעיות חשודות אשר יתפתחו באופן שונה לאחר מכן במהלך הצילוב. אם אכן נכנסת אקראיות לחישובים השונים, יתקבל גם צילוב אקראי (באופן דומה למה שקורה במציאות) וזה לא יפריע לקבל בסוף תא יחידה מצולב כפי שאנו רוצים (הרי כרגע אין לנו העדפה/בחירה לגבי אילו רביעיות יצטלבו, העיקר שיצטלבו...). **כלומר - אולי המצב שאת רואה לא מהווה בעיה כלל (היה נחמד להבין למה זה כך או כך).**

מה שלא ברור לי זה למה יש הבדל בין סריאלי ועם thread יחיד - זה לא אמור להיות אותו חישוב בדיוק? אלא אם כאמור כשהחישוב מוגדר מקבילי משהו בכל זאת (אקראי) משפיע שם.

- אני חושבת שכדאי להמשיך הלאה, בהנחה שהאקראיות איכשהו נכנסת בשל ההבדלים בין חישוב מקבילי וסריאלי וללא קשר להכנסת הפוטנציאל המאלץ.
- בדיקת שפיות - כדאי לוודא שהרביעיות שזוהו כחשודות בכל אחד מהחישובים הן רביעיות "חוקיות" (כלומר מספרי האטומים מתאימים לאטומי פחמן, חנקן, חמצן, מימן הנכונים במולקולות EPON ו-DETD).
• תריצי ריצות ארוכות כדי לאפשר במהלך כל אחת מהריצות הרבה צילובים פר תא יחידה (תזכרי שהרבה מהצילובים מתפרקים ולא נותרים יציבים, לכן צריך ריצות ארוכות עם הרבה מחזורי צילוב אפשריים כדי שחלק יתממש וייתיצב).
- מזכירה שבמאמר מאלצים לא רביעיות או שלישיות (מוותרים על האילוץ של הקשר N-H). היה מעניין להשוות את ההתנהגות עבור אחד מתאי היחידה שאת מריצה (נגיד 16x32), כלומר לראות מה ההבדל בתוצר הצילוב המתקבל עבור שני סוגי האילוץ (מעניין % צילוב מוצלח לריצות באורך אחיד).

בהצלחה,
נעמי

מאת: Yehuda Hassin
אל: Ofek Barazani
עותק: Naomi Rom
תאריך: 10 ביוני 2019, 17:52
נושא: Re: בדיקת תקינות קוד מקבילי

שלום,

אם מדובר ברביעיות חוקיות אז בודאי שאין בעיה כלל.
לא יודע איך מיקבלת את הקוד של חיפוש רביעיות, אבל כל תהליכון עובד על אטומים אחרים ולכן ימצא רביעיות אחרות וזה הגיוני.
הסדר בין התהליכונים הוא אקראי לצורך העניין...
את יכולה להשוות בין רביעיה אחת לכל תהליכון או רביעיה אחת בלבד של תהליכון אקראי,
בכל מקרה ריצות ארוכות יראו לנו אם קורה משהו...

כמובן, אני חוזר לתחילת המייל ולדברי נעמי, יש לבדוק שהרביעיות המאותרות הן חוקיות לוודא שאין פה באג יותר משמעותי. האמת, שאני קצת מופתע מההבדל בין הריצות של המקבילי לסריאלי...הייתי מצפה לזהות כמעט מוחלטת... אפשר גם לבדוק מיקומים אחרי מס' קטן של צעדי זמן של כל המערכת מעניין איך ההבדל מתפתח,

יהודה

מאת: Ofek Barazani

אל: Yehuda Hassin

תאריך: 11 ביוני 2019, 7:11

נושא: Re: בדיקת תקינות קוד מקבילי

היי

לגבי בדיקה האם הרביעיות שזוהו כחשודות חוקיות- בדקתי גם ידנית וכולן תקינות. באופן כללי הקוד בודק פעמיים- גם אם האטום הספציפי משתייך לסוג הרצוי, וגם לפי מבנה המולקולות האם הוא מתאים לרביעייה פוטנציאלית (נגיד זוגות C-O מתאימים) בכל אופן כל הרביעיות שנמצאו כחשודות בכל סוג ריצה שונה לגמרי מהרביעיות שאותרו בסוגי הריצה האחרים ולכן לא נבחרה אותה רביעייה בכל הריצות לאחר איתור כל הרביעיות העומדות בתנאים. כשהשוואתי בין מיקומים של אטומים בריצות שונות לאורך זמן הבחנתי שכלל שמתקדמים בצעדי זמן כך השגיאה במיקומים באותו אטום בין שתי ריצות שונות גדולה יותר.

אני אתחיל לבצע ריצות ארוכות היום (אתמול הסטודנטים מחומרים השתמשו במחשב במעבדה), אשתמש בקוד פייתון שכתבתי כפי שנעמי ביקשה להראות מספר מופעים של כל מולקולה לאורך הריצה ואוסיף חישוב אחוז צילוב. אבצע גם השוואה בין ריצה כפי שהיא עכשיו לריצה בה הפוטנציאל מופעל על 4 זוגות אטומים (השארתי את זה כאופציה פתוחה בקוד כל עוד מזינים פרמטרים המתאימים לזוג הרביעי)

אופק