

תרגיל בית מספר 3

נושא: סימולציה של וירוסים בעולם מונחה עצמים**דדליין: יום ה', 29/12/2022, בשעה 23:59****הגשה ביחידים****בהצלחה!****תיאור התרגיל**

בתרגיל זה תממשו עולם מונחה עצמים המוגדר עבור אוכלוסייה בגודל קבוע של אובייקטים הנקראים וירוסים, אשר מוגדרים באמצעות ווקטורים של מספרים שלמים המתארים תמורות, ותכונותיהם יתוארו בהרחבה בהמשך. מטרת התרגיל היא משולשת – ראשית, אפיון מונחה עצמים של בעיה נתונה, שנית, תרגול תכנותי של מחלקות, בנאים/מפרקים, מתודות וכן העמסת פונקציות/אופרטורים, ושלישית, מימוש של סימולציה תלוית-זמן.

וירוסים: אפיון

וירוס הינו אובייקט בעולם הסימולציה אשר מאופיין באמצעות שמו (מחרוזת תווים אלפביתיים שאורכה לכל היותר 8) וכן באמצעות ווקטור שלמים באורך קבוע ℓ שמגדיר **תמורה**. ההנחה היא כי בעת בדיקת מצב העולם (כלומר בכל פעימת זמן בדירה, או דור) קיימת אוכלוסייה בגודל נתון וקבוע של וירוסים, כל וירוס מכיל תמורה חוקית, וכי הם מבצעים עדכוני ווקטורים בין נקודת זמן אחת לאחרת באופן שיפורט.

עם הגדרת העולם, יוגדר גם **ווקטור מטרה**, אליו שואפת להגיע אוכלוסיית הווירוסים. בכל נקודת זמן ניתן לחשב את מספר הפגיעות של כל וירוס ביחס לווקטור המטרה (כלומר, מספר הקואורדינטות בווקטור הווירוס בהן המספר זהה למספר המטרה באותו מיקום), ומכאן ניתן להגדיר את מידת השגיאה של וירוס נתון ביחס למטרה. הווירוס הדומה ביותר לווקטור המטרה, כלומר בעל מידת השגיאה הנמוכה ביותר, ייחשב לווירוס הטוב באותה נקודת זמן (*fittest*). בכל מעבר דור הסימולציה תעדכן את זהות הווירוס הטוב ביותר בהיסטוריה שלה (במקרה שנדרש עדכון).

מידת השגיאה של וירוס נתון ביחס לווקטור המטרה

יהי מימד וקטור השלמים של הווירוסים ℓ , ותהי \vec{a} מחרוזת חוקית של וירוס נתון המקיימת

$$\vec{a} \in \pi\{1 \dots \ell\} = \begin{pmatrix} 1 & 2 & 3 & \dots & \ell \\ a_1 & a_2 & a_3 & \dots & a_\ell \end{pmatrix}$$

תהי $\vec{\tau}$ מחרוזת המטרה המקיימת גם-כן $\vec{\tau} \in \pi\{1 \dots \ell\}$. נגדיר את מספר הפגיעות של הווקטור \vec{a} בווקטור המטרה $\vec{\tau}$ כמספר הקואורדינטות בהן מתקיים שוויון:

$$h(\vec{a}, \vec{\tau}) = |\{k \mid a_k = \tau_k, k = 1 \dots \ell\}|$$

בהתאם, **מידת השגיאה** ε של הווירוס הנושא את הווקטור \vec{a} ביחס לווקטור המטרה $\vec{\tau}$ הינה סקלר ב-[0,1] המוגדר כהפרש מספר הפגיעות המנומלל מהיחידה:

$$\varepsilon_{\vec{\tau}}(\vec{a}) = 1 - \frac{h(\vec{a}, \vec{\tau})}{\ell} \in [0,1]$$

אחד מתנאי העצירה של הסימולציה הוא הימצאותו של ווקטור בעל מידת שגיאה 0.0 אצל אחד מהווירוסים באוכלוסייה. כמו-כן, התכנית צריכה לעקוב אחר מידת ההתאמה של הווירוסים בכל דור, וכאמור, לשמור בזיכרון את הווירוס "הטוב ביותר" בכל הדורות (ווקטור התמורה ומידת השגיאה של הווירוס הטוב ביותר בכל היסטוריית הריצה).

עדכון הווירוסים מתרחש בכל פעימת זמן בדידה (דור), והוא **הסתברותי** בטבעו. קיימים שני סוגי עדכון – עדכון קבוצתי, ועדכון פרטני – כפי שיוסבר בפירוט בסעיף זה.

עדכון קבוצתי: בכל דור ייזרק מן האוכלוסייה הווירוס בעל הווקטור עם השגיאה הגדולה ביותר, ויוחלף בעותק של הווירוס בעל הווקטור עם השגיאה הנמוכה ביותר באותו דור. שם הווירוס המוחלף יהיה ווריאציה של שם הווירוס **המחליף** – שם הווירוס המקורי, קו תחתי, ואינדקס המונה את מספר הווריאציה מאז תחילת הסימולציה (המחשה תינתן בהמשך ההנחיות).

עדכון פרטני: בהינתן מדד התמורות $p_m \in \{1 \dots \ell\}$ (מספר חיובי שנתון כפרמטר של הסימולציה), הווקטור יעבור החלפות של צמד קואורדינטות במיקומים אקראיים לפי מדד זה (כלומר, מספר ההחלפות יהיה כמדד התמורות, והצמדים ייבחרו באקראיות). לצורך המחשה, כך יתרחש עדכון פרטני במעבר דור בהינתן מדד תמורות $p_m = 1$ (צמד יחיד, $\{i, j\}$, נבחר באקראיות ומוחלף):

$$\vec{a}^{(g)} = \begin{pmatrix} 1 & \dots & 2 & \dots & 3 & \dots & \ell \\ a_1 & \dots & a_i & \dots & a_j & \dots & a_\ell \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \dots & 2 & \dots & 3 & \dots & \ell \\ a_1 & \dots & a_j & \dots & a_i & \dots & a_\ell \end{pmatrix} = \vec{a}^{(g+1)}$$

העדכון הפרטני מתרחש בכל דור עבור כל הווירוסים (כולל הווירוס שהוחלף בעדכון הקבוצתי); מדד התמורות הינו לפחות 1, כך שכל פרטי האוכלוסייה בהכרח יעודכנו במעבר מדור לדור.

הרעיון הכללי של התכנית אותה תממשו הוא לאתחל אוכלוסיית וירוסים לפי קלט מן המשתמש, ולסמלץ את תוכן הווקטורים (תמורות), במעברי זמן בדידים, עד להגעה של הווירוס הראשון לווקטור המטרה (מידת שגיאה 0.0), אז תסתיים הסימולציה. בכל מעבר זמן (דור), יעודכנו הווקטורים של הווירוסים לפי מתכונת העדכון ההסתברותית שלהם. בסופה של התכנית, הווקטורים של אוכלוסיית הדור האחרון יימסרו כפלט למשתמש, בנוסף לווקטור הווירוס הטוב ביותר בהיסטוריית הריצה.

פורמט קלט

התכנית תקבל כקלט שני קבצים – הראשון יתאר את עולם הווירוסים (פרמטרים שונים, וקטור המטרה וכיו"ב), והשני יכלול את הדור ההתחלתי (יש לבדוק כי הפרמטרים אכן הועברו במידה ולא על התוכנית להדפיס הודעת שגיאה

Usage: run the program with <init file name> <location file name>

לאחר קומפילציה של התכנית לכדי קובץ הרצה בשם `simVirus`, כך ניתן יהיה להריץ אותה במקרה של עבודה דרך טרמינל:

```
>> ./simVirus config.dat first_generation.dat
```

1. מרכיבי עולם הווירוסים, אשר יוזנו לתכנית בתחילתה בקובץ הראשון (`config.dat` בדוגמא שלעיל), יתוארו בפורמט הבא: השורה הראשונה תכיל את מימד ווקטורי התמורות ℓ , השורה השנייה תכיל את מדד התמורות p_m עבור העדכון הפרטני, והשלישית תכיל את ווקטור המטרה \vec{c} (באמצעות ℓ קואורדינטות מופרדות ברווחים). להלן דוגמא של קובץ קלט חוקי של תיאור עולם הווירוסים:

```
10
1
5 2 3 10 6 9 8 7 1 4
```

2. על קובץ הקלט השני (קובץ הדור ההתחלתי; `first_generation.dat` בדוגמא שלעיל), להכיל בשורה הראשונה את מספר הווירוסים בעולם (גודל האוכלוסייה שיישאר קבוע במהלך הריצה), ובשורות לאחר מכן את שמם של כל הווירוסים ומצבם ההתחלתי. כל שורה תתחיל בשמו של הווירוס, ולאחר מכן תכיל את הווקטור ההתחלתי שלו, באמצעות ℓ קואורדינטות מופרדות ברווחים. להלן דוגמא של קובץ קלט חוקי של אתחול התאים:

```
5
csMEX 1 2 3 4 5 6 7 8 9 10
ukSFO 10 2 3 4 5 6 7 8 9 1
fhOAK 3 1 2 4 5 6 7 10 9 8
```

תכנות בשפת ++C, סתיו 2022-23

jkFIN 7 2 3 4 5 1 6 8 10 9
mjdjAMS 8 2 3 9 5 6 10 7 1 4

3. במקרה של קלט לא חוקי (למשל, שגיאת קבצים, חוסר-עקביות במימדים וכיוצא באלה), התכנית תסיים את ריצתה בשלב זה ותדפיס הודעת שגיאה מתאימה לערוץ השגיאות הסטנדרטי. ("Invalid Input\n")
4. בהנחה שהקלט של המשתמש חוקי, המשתמש יישאל בשלב זה למספר מעברי הזמן המקסימליים אותם ירצה שהתכנית תסמלץ (שאלה/תשובה באמצעות פלט/קלט סטנדרטיים, בהתאמה). בעקבות כך על התכנית להתחיל את מעברי הזמן בעולם, תוך עדכון הווקטורים של הוורוסים מדור לדור. התכנית תרוץ עד שאחד הוורוסים יגיע למידת שגיאה 0.0, או עד שמספר המעברים המקסימליים יחלוף.

פורמט פלט

בסיום מעברי הזמן על התכנית להחזיר כפלט את ווקטורי התמורות של הדור האחרון. על נתונים אלו להיות מודפסים דרך ערוץ הפלט הסטנדרטי. בתום הדפסת האוכלוסייה, תודפס שורת רווח, ולאחריה הוורוס הטוב ביותר בריצה. למשל, כך יכול להיראות הפלט עבור הדוגמא שלעיל (שימו לב להשתלטות של וורוס מסוים על האוכלוסייה) –

```
mjdjAMS_112 4 2 3 5 9 6 10 7 1 8  
mjdjAMS_666 8 2 3 9 6 5 10 7 1 4  
mjdjAMS_45 1 2 3 9 5 6 4 7 8 10  
mjdjAMS_786 9 2 3 8 4 6 10 7 1 5  
mjdjAMS_987 8 10 3 9 5 6 4 7 1 2  
  
mjdjAMS_341 5 2 3 10 9 6 8 7 1 4
```

דרישות מימוש

אתם חופשיים לתכנן ולממש את התכנית כרצונכם, פרט לדרישות הבאות אותן עליכם לקיים במימוש שלכם:

- עבור אובייקט הוורוס יש לממש באופן מפורש את ה-**Big Five**, גם אם מימוש ברירת-מחדל של הקומפיילר עשוי להספיק מבחינת פונקציונאליות. בשאר המחלקות אין חובה לכתוב את חמשת הגדולים, אך נדרש לתת התייחסות במידה ובחרתם לא לממש.
- עבור אובייקט האוכלוסייה, יש לממש אופרטור ++ (postfix) המבצע עדכון קבוצתי.
- עבור אובייקט הוורוס, יש לממש את האופרטור * (unary) המבצע עדכון פרטי עבור הווקטור של הוורוס.

הנחות עבודה

- שם התחלתי של וורוס הוא מחרוזת תווים אלפביתיים באורך 8 לכל היותר.
- ווריאציה של שם וורוס (בעת העדכון הקבוצתי) היא שרשרת של שם הוורוס המחליף המקורי עם קו תחתי והאינדקס המונה את מספר הווריאציה מאז תחילת הסימולציה.
- מספר מעברי הזמן (דורות סימולציה) הינו מספר טבעי: בהינתן ערך 0, התכנית תדפיס את מצב העולם ההתחלתי ותסיים; בהינתן ערך 1, התכנית תחשב מעבר דור בודד, תדפיס את ווקטורי הוורוסים ותסיים; וכך הלאה. ניתן להניח כי מספר הדורות לא יעלה על 10^6 .
- גודל האוכלוסייה הוא לפחות 2 וורוסים; ניתן להניח כי יש לכל היותר 10,000 וורוסים.
- מימד ווקטורי התמורה הוא לפחות 5; ניתן להניח כי הממד לא יעלה על 100.

דגשים

- יש לתכנן מראש את מבנה התכנית, ולהגדיר בהתאם את האובייקטים איתם תעבדו; הקפידו על תכנון מונחה-עצמים של עולם הסימולציה על מרכיביה השונים.
- הקדישו מחשבה למבני הנתונים השונים בהם תעשו שימוש; הביאו בחשבון שיקולי יעילות.
- יש לבדוק תקינות קלטים ולהציג הודעות שגיאה מתאימות.
- אין אפשרות להשתמש בספריית STL בתרגיל זה.

תכנות בשפת ++C, סתיו 2022-23

- עליכם לוודא כי התכנית עוברת קומפילציית ++g התואמת את הקומפיילר שעל שרת המכללה ללא כל שגיאות או אזהרות כלשהן, ורצה בהצלחה.
- עליכם לתעד את הקוד באמצעות הערות המתארות בקצרה את הפונקציות השונות.
- עצה: מומלץ להשתמש בפונקציה `rand()` ובקבוע `MAX_RANDOM` לצורך הגרלה אקראית המתפלגת באופן אחיד במימוש
- אין להשתמש בפונקציה `srand()` לקביעת `seed` עבור המתודה `rand()`.
- העדכונים ההסתברותיים. הסברים מפורטים בנוגע למרכיבים ההסתברותיים יינתנו בתרגול הקרוב.

הגשה

- עליכם להגיש במערכת Moodle קובץ ארכיב מטיפוס zip בלבד, ששמו כולל את קוד הקורס ('44'), שם התרגיל ('ex3') ותעודת הזהות של הסטודנט/ית המגישה, מופרדים בקו תחתי בפורמט הבא: `44_ex3_studID.zip`.
- על ארכיב zip זה להכיל את כל קבצי המקור (ממשק/מימוש) הנדרשים לקומפילציה, והוא רשאי להכיל תיעוד טקסטואלי; מבחינת טיפוס קבצים, עליו לכלול רק קבצים עם סיומות `*.cpp`, `*.h`, `*.txt`.
- לדוגמה: על סטודנט שמספר הזיהוי שלו הינו 012345678 להגיש ארכיב בשם `44_ex3_012345678.zip` הכולל את כל קבצי המקור של הפרויקט, ללא תיקיות כלשהן, ורשאי להכיל קובץ טקסטואלי לתיעוד.

אי-הקפדה על ההנחיות, כולל פורמט ההגשה הדיגיטלי, תגרור הורדה בציון התרגיל.
לא תתקבלנה הגשות באיחור!