

דוח ביולוגיה חישובית – תרגיל 2:

קישור לgit ניתן למצוא כאן.

הוראות הרצה – מצורפים שישה קבצים לבדיקת התוכנית. כל קובץ מגדיר לוח בגודל מסוים וברמת קושי מסוימת (ניתן לראות לפי שם הקובץ).
אופציה ראשונה להרצה – הורדת קובץ הפייטון, התקנת הספריות, והרצת הפקודה
`python futoshiki.py input.txt` כך ש הוא שם קובץ הקלט ונמצא באותה תיקייה עם קובץ הפייטון.
אופציה שניה להרצה – באמצעות קובץ `exe`. הפקודה להרצה: `futoshiki.exe input.txt` כך ש הוא שם קובץ הקלט ונמצא באותה תיקייה עם קובץ `exe`.

הסבר מימוש האלגוריתם הגנטי:

- **ייצוג פתרונות** – מטריצה ריבועית, בגודל שהוגדר ($N \times N$).
- **פונקציית ההערכה** – עבור כל לוח, ספרנו את מספר האילוצים שהופרו.
הפרת אילוץ יכולה להיות שעבור זוג משבצות לא מתקיים $<$ כפי שהוגדר בקובץ הקלט.
הפרת אילוץ נוספת תהיה חזרה של מספר מסוים יותר מפעם אחת בשורה/ בעמודה.
נספור את מספר האילוצים שהופרו עבור כל לוח – זו תהיה ההערכה.
בכל דור, הלוח שיש לו את ההערכה הנמוכה ביותר, הוא הלוח הטוב ביותר מאותו דור.
לוח שעבורו ההערכה שווה ל0, מהווה פתרון מושלם לבעיה.
- **פונקציית ההכלאה** – קודם כל, נרצה לבחור זוג הורים לזיווג. נרצה לבחור את הייצורים המבטיחים ביותר, אך עם זאת, נרצה לשמור על גיוון האוכלוסייה (ולא לבחור כל הזמן את אותו זוג).
אז ביצענו את בחירת זוג ההורים מהדור הקודם כך – נתנו הסתברות גבוהה יותר להיבחר ללוחות הטובים, והסתברות נמוכה ללוחות הפחות טובים. כך גם שמרנו על גיוון האוכלוסייה וגם השתדלנו לבחור את ההורים הטובים ביותר – כדי להתקרב לפתרון. ההכלאה עצמה – בחרנו מספר אקראי בין 0 לN, נסמנו בא. את x השורות הראשונות לקחנו מההורה הראשון ואת N-x השורות הנותרות לקחנו מההורה השני.
- **מוטציות** – עבור כל לוח, סרקנו אותו, ובהסתברות של $\frac{1}{10}$ החלפנו את הערך שהיה באותו תא בערך רנדומלי.
כמובן שלאחר המוטציות סרקנו שוב את הלוח ושמנו במקומות הנדרשים את המספרים שקיבלנו כקלט במקומות שבהם הם נדרשו להיות (אם היו כאלה).
- **בעיית ההתכנסות המוקדמת** – בריצות מסוימות הגענו למצב שהאלגוריתם נתקע באופטימום מקומי, ולא מצליח למצוא את האופטימום הגלובלי.
למשל, עבור ריצות רבות הפתרון הסופי שהתקבל היה בעל `evaluation = 2`, כלומר מאוד קרוב לפתרון (רק שני אילוצים לא מסופקים).
על מנת להוציא את האלגוריתם מהמצב הזה, ולאפשר לו למצוא את האופטימום הגלובלי, כל מספר קבוע של דורות ביצענו restart לאוכלוסייה. הגרלנו אוכלוסייה חדשה, שעבורה ייתכן שנגיע לאופטימום אחר (בתקווה שיהיה הגלובלי).
באופן שקול, היה אפשר להעלות את ההסתברות למוטציות במצב הזה, ולקבל השפעה דומה.

רני גוטליב 322805029
הדס בביוב 322807629

- **בחירת הדור הבא** – בחרנו לממש אליטיזם. בכל דור בחרנו את 4 הייצורים הטובים ביותר (עם ההערכה הכי נמוכה), והעברנו אותם לדור הבא. את שאר האוכלוסייה (96 הייצורים הנותרים), ייצרנו בעזרת פונקציית ההכלאה, כפי שהסברנו למעלה.
- **מספר דורות** – שמנו לב שככל שמדובר בלוח גדול יותר, לוקח לאלגוריתם יותר דורות להשתפר ולהתכנס לפתרון. אז החלטנו שמספר הדורות יהיה כפונקציה שתלויה ב-N (קבוע כפול N).

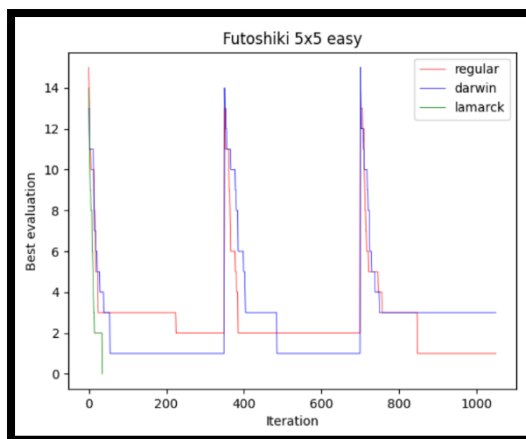
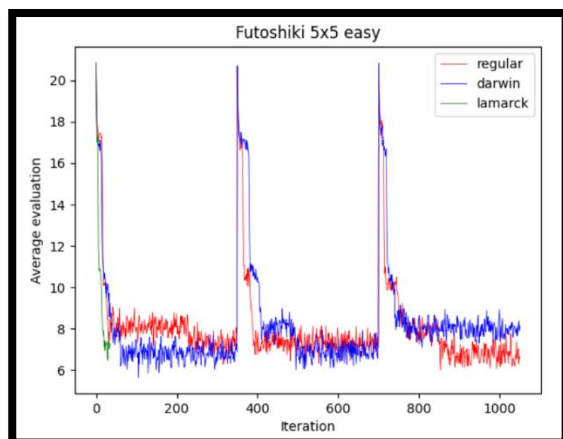
נוסיף, שהחזקנו משתנה ששמר את הפתרון הכי טוב שראינו עד כה (כלומר, בעל ההערכה המינימלית). החזרנו אותו בסוף הריצה במקרה שלא מצאנו את הפתרון המושלם.

השוואה בין שלושת האלגוריתמים:

אינטואיטיבית, האלגוריתמים שמשתמשים באופטימיזציות, אמורים להגיע לפתרון עם הערכה טובה יותר מהאלגוריתם הרגיל – שלא מבצע אופטימיזציות.

נערוך השוואה בין שלושת האלגוריתמים על לוחות בשלושה גדלים שונים (5x5, 6x6, 7x7), בשתי רמות קושי שונות (easy, tricky).

נתחיל מלוח ראשון בגודל 5x5 וברמת קושי easy:



כפי שניתן לראות, האלגוריתם הלאמארקי מצא פתרון בשלבים מוקדמים (מצא לוח שקיבל הערכה 0 והחזיר אותו), ולכן העקומה הירוקה מסתיימת בשלב מוקדם.

הפתרון שהודפס הוא:

```
[[1. 5. 4. 2. 3.]
 [2. 4. 3. 1. 5.]
 [3. 2. 1. 5. 4.]
 [4. 1. 5. 3. 2.]
 [5. 3. 2. 4. 1.]]
```



שתואם לפתרון שבאתר:

רני גוטליב 322805029
הדס בביוב 322807629

ניתן לראות שמדובר במגמת ירידה (כלומר ההערכה משתפרת עם הזמן) ולאחר מספר קבוע של איטרציות יש גל של עליה שמסמל את הrestart. בנקודה הזו מתחילים מלוחות רנדומליים חדשים ולכן ההערכה חוזרת להיות גבוהה.

בנוסף, בגרף הימני ניתן לראות שההערכה המינימלית תמיד משתפרת (או נשארת קבועה) זאת בגלל האליטיזם – שדאגנו להעביר לדור הבא את הייצורים הכי טובים. לעומת זאת, בגרף השמאלי נקבל מגמה כוללת של ירידה אך יתכנו תנודות קלות. האלגוריתם הרגיל הגיע לפתרון עם הפרה של אילוץ אחד, נצרף תמונה:

[1. 3. 4. 5. 2.]
[2. 4. 1. 3. 5.]
[3. 2. 5. 1. 4.]
[4. 5. 3. 2. 1.]
[5. 1. 2. 4. 3.]

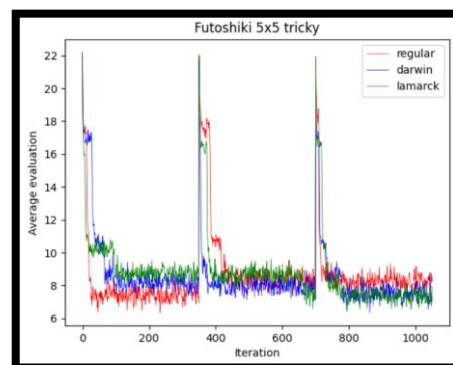
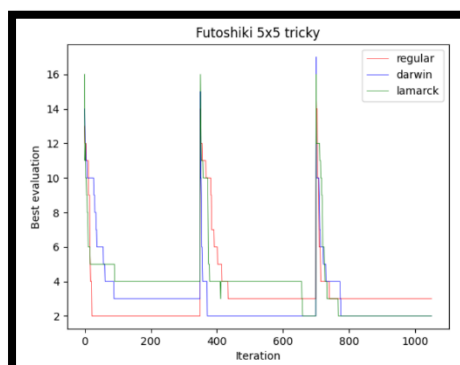
האילוץ שהופר – ה2 שבשורה השלישית לא יותר גדול מ5 שבשורה הרביעית. והאלגוריתם של דארווין הגיע לפתרון עם הפרה של שלושה אילוצים:

[1. 5. 2. 3. 4.]
[5. 1. 3. 4. 5.]
[2. 4. 5. 1. 3.]
[3. 2. 4. 2. 1.]
[4. 3. 1. 5. 2.]

האילוצים שהופרו פה:

1. בשורה השנייה הספרה 5 מופיעה פעמיים.
2. בשורה הרביעית הספרה 2 מופיעה פעמיים.
3. האילוץ גדול מ, שסומן באדום (2 לא גדול מ4).

נמשיך ללוח השני, בגודל 5x5 וברמת קושי tricky:

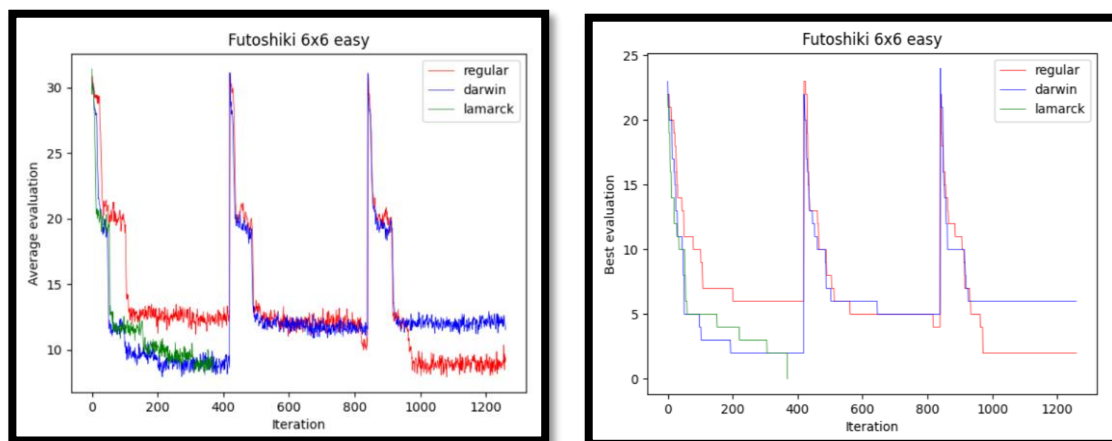


בגרף מימין, ניתן לראות שרוב האיטרציות האלגוריתמים עם האופטימיזציות, בעלי ממוצע טוב יותר (נמוך יותר).

רני גוטליב 322805029
הדס בביוב 322807629

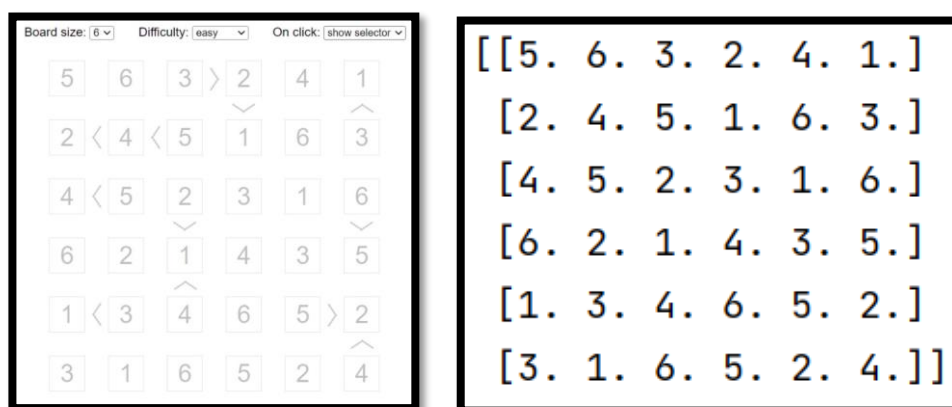
וכפי שניתן לראות בגרף השמאלי, שני אלגוריתמי האופטימיזציות (דארווין והלאמארקי) החזירו פתרון עם הערכה 2, והאלגוריתם הרגיל החזיר פתרון עם הערכה 3 – שהיא קצת פחות טובה ביחס לאלגוריתמים האחרים.

כעת, נעבור ללוחות גדולים יותר, בגודל 6x6. נתחיל מרמת קושי easy:



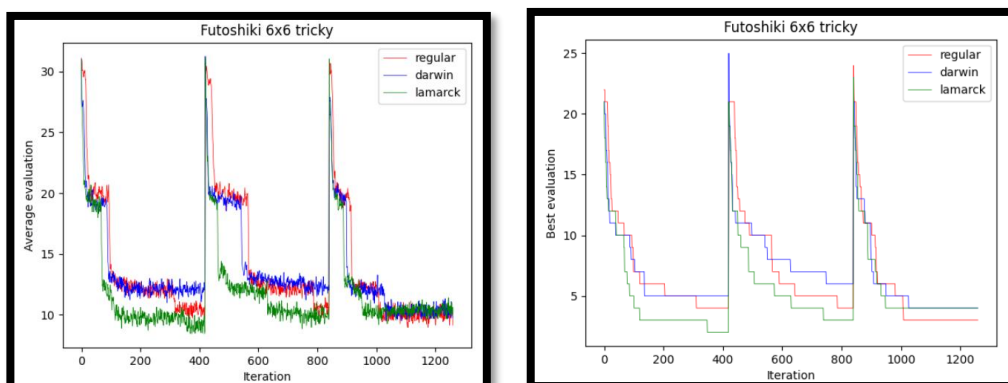
גם במקרה הזה, האלגוריתם הלאמארקי מצא פתרון מושלם עוד לפני הrestart הראשון. האלגוריתם הדארוויני מצא פתרון עם הערכה 2 (כפי שניתן לראות בגרף הימני, הוא מחזיק בפתרון הנ"ל באיזור איטרציה 200 עד לrestart הראשון). זה הפתרון שמוחזר כיוון שלאורך כל הדרך אנו נעדיכן במשתנה את הפתרון הכי טוב שמצאנו, גם אם אפסנו את האוכלוסייה ע"י restart. כלומר, מוחזר הפתרון בעל ההערכה המינימלית שמצאנו לאורך כל האיטרציות. האלגוריתם הרגיל מצא פתרון עם הערכה 2 – כמו האלגוריתם הדארוויני.

נצרך את הפתרון המושלם שמצא האלגוריתם הלאמארקי ואת הלוח מהאתר:



רני גוטליב 322805029
הדס בביוב 322807629

נמשיך ללוח בגודל 6x6, רמת קושי tricky:

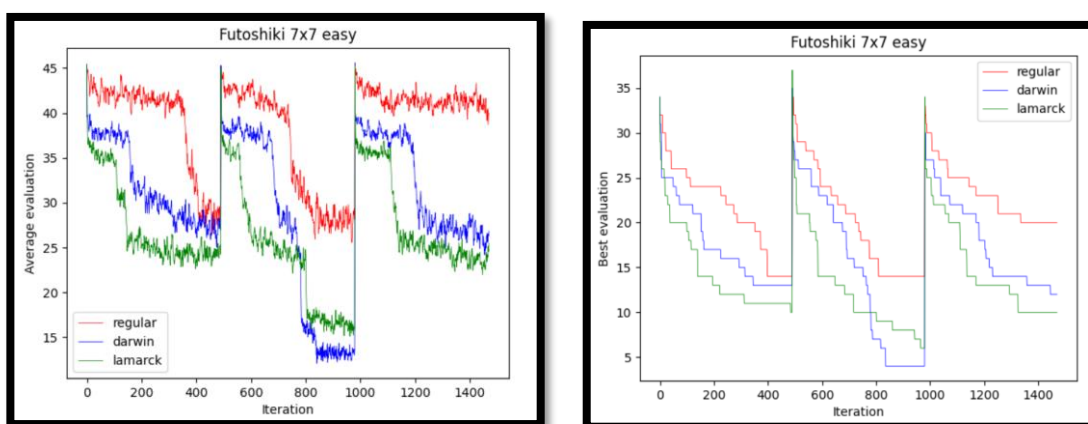


בגרף השמאלי ניתן לראות שהאלגוריתם הלאמארקי עם ממוצע נמוך לאורך כל הדרך ולקראת הסוף האלגוריתמים האחרים מתמזגים איתו.

מבחינת הפתרון הסופי, האלגוריתם הלאמארקי החזיר פתרון עם הערכה 2, בעוד האלגוריתם הרגיל החזיר לוח בעל הערכה 3 והאלגוריתם הדארוויני החזיר לוח בעל הערכה 4.

נעבור לגודל האחרון 7x7.

רמת קושי easy:



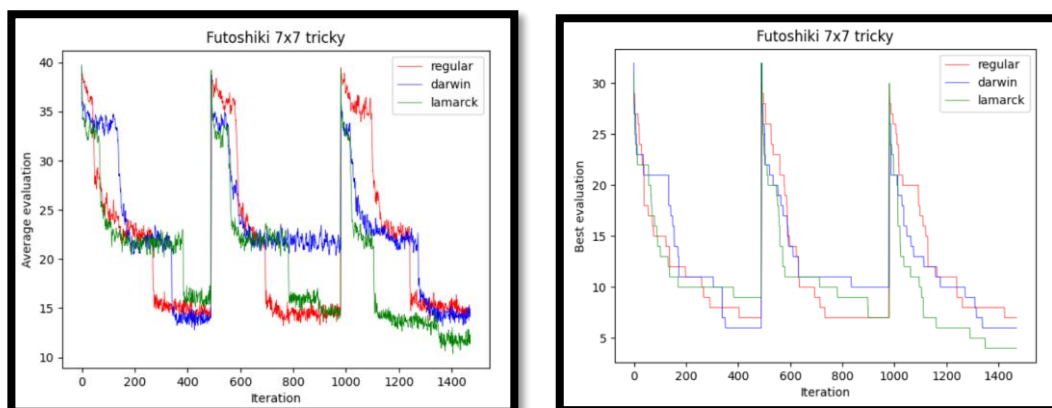
בלוח הזה, כפי שניתן לראות, האלגוריתם הדארוויני החזיר את הפתרון הטוב ביותר עם הערכה 4 (הקטנה ביותר).

האלגוריתם הלאמארקי החזיר פתרון קרוב לדארוויני עם הערכה 6, והאלגוריתם הרגיל החזיר פתרון עם הערכה 14 – שזה הלוח הטוב ביותר שהוא הצליח למצוא במשך כל האיטרציות.

בגרף השמאלי, ניתן לראות שהממוצע של האלגוריתם הרגיל, מאוד גבוה לאורך כל התהליך. ולכן לא מפתיע שקיבלנו פתרון סופי עם הערכה גבוהה.

רני גוטליב 322805029
הדס בביוב 322807629

נמשיך לרמת הקושי האחרונה על לוח tricky, 7x7:



התוצאות שהתקבלו פה:

לאמארקי – 4

דארוויני -6

רגיל – 7

מסקנות:

- כפי שחשבנו לפני, האלגוריתמים עם האופטימיזציות נתנו תוצאות טובות יותר בדו"כ, והחזירו פתרונות שמקיימות יותר אילוצים מהאלגוריתם הרגיל.
- נשים לב, שיש גורם נוסף שמשפיע על התוצאות. האוכלוסייה הראשונית שמוגרלת, וגם האוכלוסיות שמוגרלות אחרי restart-ים הן רנדומליות לחלוטין. ייתכן שאוכלוסייה מסוימת מכילה לוח שמאוד קרוב לפתרון, זה יכול לשפר את התוצאות שלו.