מעבדה בבינה מלאכותית

דו"ח תרגיל בית 1 – Genetic Algorithms

שמות:

אסיל נחאס, 212245096

עוביידה חטיב, 201278066

**הרצת הקוד**

הקוד ניתן להרצה דרך שורת הפקודות תוך קבלת ארגומנט של מגבלת זמן הריצה בשניות. לדוגמה:

python lab1.py 60

**ייצוג הפרטים והאוכלוסייה**

בקוד שלנו, הפרטים מיוצגים באמצעות מופעים של המחלקה Individual, השומרת עבור כל פרט את הגינום ואת הפיטניס שלו. כמו כן, המחלקה מכילה מתודה המחשבת ומעדכנת את משתנה הפיטניס של הפרט.

האוכלוסייה מיוצגת ע"י מופע של המחלקה Population, המכילה משתנים עבור גודל האוכלוסייה, מחרוזת המטרה, ורשימת הפרטים של האוכלוסייה. בנוסף, המחלקה מכילה את המתודות הבאות: מתודה המאתחלת אוכלוסייה כקבוצה של פרטים עם גנומים אקראיים באורך מחרוזת המטרה, מתודה המחשבת ומעדכנת את הפיטניס של כל אחד מפרטי האוכלוסייה, המתודה הממיינת את פרטי האוכלוסייה לפי ערכי הפיטניס שלהם, ומתודה המעבירה את הפרטים עם הפיטניס הכי טוב לדור הבא (Elitism).

להלן המבנים ההתחלתיים של שתי המחלקות, אליהם יתווספו עוד משתנים ומתודות לפי הנדרש במעבדה.

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

**בחירת הפרמטרים**

**תנאי העצירה בשל התכנסות מקומית** נבחר להיות כאשר הפיטניס הטוב ביותר לא משתפר במשך 50 איטרציות. ההחלטה מתבססת על תצפיות שלנו שהראו כי אם לא מתרחש שיפור בתוך 50 דורות, אז לא יתרחש גם לאחר 100,200 ואפילו 400 איטרציות נוספות. ההחלטה אודות תנאי העצירה התקבלה עוד לפני בחירת מספר הפרטים באוכלוסייה, ולכן נלקחו בחשבון כל גדלי האוכלוסייה הפוטנציאלים, תוך ביצוע 100 הרצאות עבור כל אחד מהגדלים.

תנאי עצירה נוספים הם התכנסות גלובלית והגעה לסף הזמן הנתון כקלט. התכנסות גלובלית מתרחשת כאשר ערך הפיטניס של הפתרון הטוב ביותר בדור כלשהו שווה ל- 0. הגעה לסף הזמן נבדקת בסיום כל דור, ואם הזמן שחלף מאז תחילת הריצה הוא יותר מזה שניתן כארגומנט האלגוריתם נעצר.

תמונה שמכילה טקסט, גופן, צילום מסך, מספר

התיאור נוצר באופן אוטומטי

**המספר המקסימלי של דורות** הוחלט להיות 120. ערך זה גדול מכל מספר דורות מקסימלי שהתקבל ב- 400 הרצות שהתבצעו כמתואר מקודם, ולכן הוא מבטיח באופן כמעט וודאי שהאלגוריתם לא ייעצר בשל מספר לא מספיק של דורות.

תמונה שמכילה טקסט, גופן, לבן, צילום מסך

התיאור נוצר באופן אוטומטי

בחירת **גודל האוכלוסייה** נעשתה על סמך ניתוח סטטיסטי שנעשה על 100 הרצות לכל גודל אוכלוסייה פוטנציאלי. המטרה הייתה לבחור את האוכלוסייה הקטנה ביותר המתכנסת גלובלית בלפחות 95% מהמקרים. הגודל הראשון שענה על הקריטריון הוא 8192 ועל כן נבחר.

תמונה שמכילה טקסט, גופן, קבלה, לבן

התיאור נוצר באופן אוטומטי

סעיף 1

הוספנו למחלקה Population מתודה המחשבת ומדפיסה את הפרט עם ה- Fitness הגבוה ביותר, הפרט עם ה- Fitness הנמוך ביותר, ממוצע, סטיית תקן, וטווח ה- Fitness של האוכלוסייה.

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

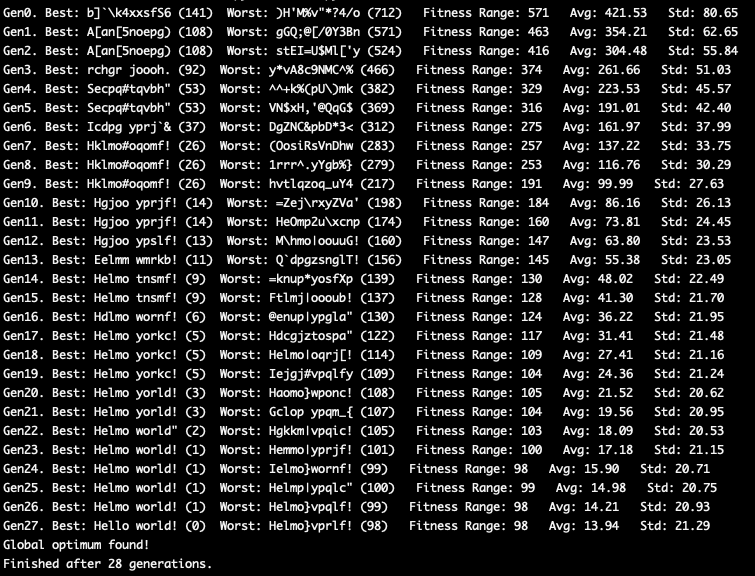
מצורף תוצאות של 3 הרצות שנעשו:

תמונה שמכילה צילום מסך, דפוס, מונוכרום, שחור ולבן

התיאור נוצר באופן אוטומטי

תמונה שמכילה מונוכרום, טקסט, צילום מסך, שחור ולבן

התיאור נוצר באופן אוטומטי



סעיף 2

חישוב והדפסת שני זמני הריצה, ה- Clock tick והאבסולוטי התבצעה ע"י אתחול שני משתנים, המחזיקים את הזמן של תחילת הביצוע :

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

בנוסף, את הפונקציה הבאה המחשבת ומדפיסה את הזמן שעבר מאז תחילת ההרצה:

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

להלן המספרים שהתקבלו מהרצה אחת:

תמונה שמכילה טקסט, צילום מסך, שחור ולבן, דפוס

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט, צילום מסך, שחור ולבן, גופן

התיאור נוצר באופן אוטומטי

סעיף 3

שני הגרפים מומשו באמצעות שתי הפונקציות הבאות שהתווספו למחלקת Population, ואשר נקראות בתום ריצת האלגוריתם:

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

מצורף דוגמה לפלט שהתקבל בעקבות אחת ההרצות. ה- Boxplots, בשל היותם 27, צורפו חלקית, לסירוגין. בנוסף, לכל סוג של גרפים מצורף הסבר אודות מה הוא מבטא.

תמונה שמכילה קו, עלילה, תרשים

התיאור נוצר באופן אוטומטי

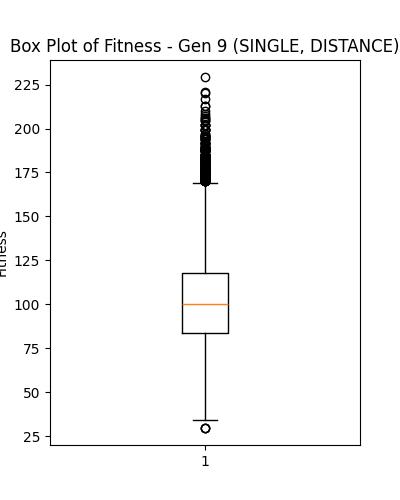
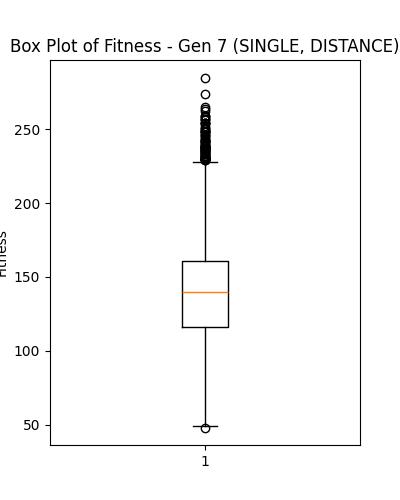
הגרף מציג את התקדמות האלגוריתם לאורך הדורות משלושה נקודות מבט של הפיטנס הטוב ביותר, הגרוע ביותר, והממוצע. ניתן לראות שכולם משתפרים ככל שהדור יותר גבוה, במיוחד של הפיטנס הטוב ביותר אשר מעיד על התכנסות האלגוריתם לפתרון. כמו כן, מהגרף ניתן להפיק ידע לגבי טווח הפיטנס שנהיה יותר קטן לאורך הדורות.

תמונה שמכילה טקסט, צילום מסך, תרשים, קו

התיאור נוצר באופן אוטומטיתמונה שמכילה טקסט, צילום מסך, תרשים, קו

התיאור נוצר באופן אוטומטיתמונה שמכילה טקסט, תרשים, צילום מסך, קו

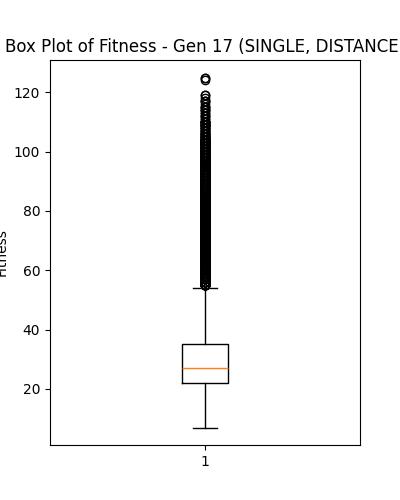
התיאור נוצר באופן אוטומטי

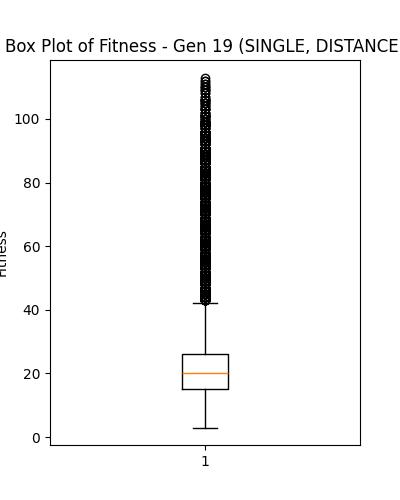
תמונה שמכילה טקסט, צילום מסך, תרשים

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט, צילום מסך, תרשים, קו

התיאור נוצר באופן אוטומטיתמונה שמכילה טקסט, צילום מסך, תרשים, קו

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט, צילום מסך, תרשים, מלבן

התיאור נוצר באופן אוטומטיתמונה שמכילה טקסט, צילום מסך, מלבן, תרשים

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט, צילום מסך, מלבן, תרשים

התיאור נוצר באופן אוטומטי

הגרף מאפשר להבין את פיזור ערכי הפיטניס של האוכלוסייה בכל דור, ומסייע להבחין בין אם האוכלוסיה מתכנסת לפתרון או נתקעת על ערכים (מידת ה- Exploration לעומת ה- Exploitation). ניתן לראות בדוגמה שהערכים מתכנסים לפתרון הטוב ביותר עם הזמן. כמו כן, גם הגרף הזה נותן מדד אודות מידת פיזור הערכים, ובגרף רואים שהשונות של הערכים קטנה עם הזמן.

סעיף 4

התווספו שלושת הפונקציות המצורפות למטה, כל אחת מבצעת סוג של שחלוף (Crossover). אחת מהן נקראת לפי אופרטור השיחלוף הנבחר.

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט, צילום מסך, גופן

התיאור נוצר באופן אוטומטי

סעיף 5

חלקים באלגוריתם האחראים ל- Exploration:

* מוטציות: מאפשרות אקראיות וחשיפה לגנומים חדשים ולא צפויים.
* שיטות שחלוף גבוהות, במיוחד Uniform, אשר יוצרת צאצאים מורכבים ושונות בהשוואה להורים שלהם.
* בחירת הורים ממרחב בגודל חצי האוכלוסייה, שהוא מרחב גדול יחסית, כך שמאפשר גיוון ומונע חמדנות.

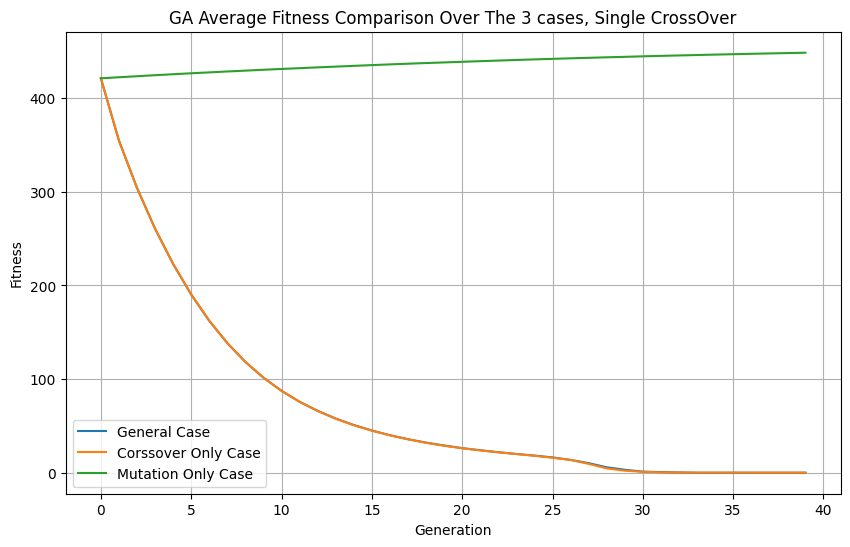
חלקים באלגוריתם האחראים ל- Exploitation:

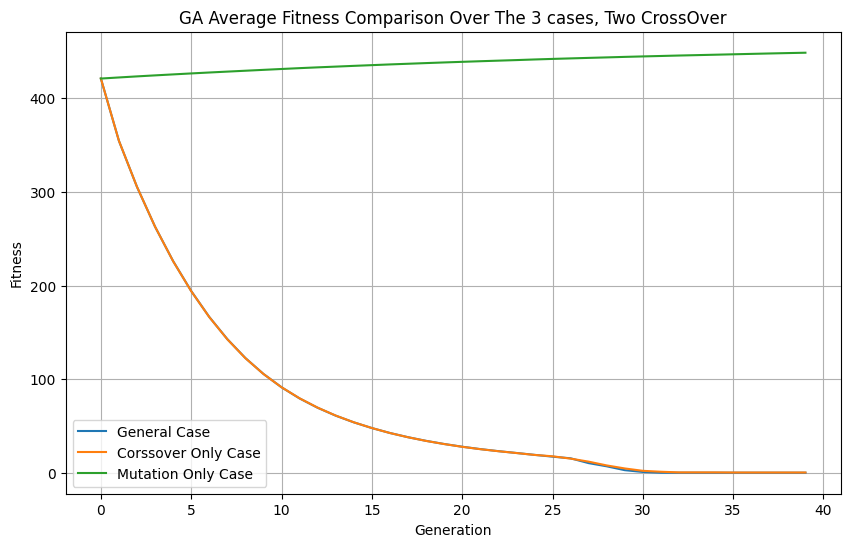
* אליטיזם: בכך שהוא שומר את הפתרונות הטובים ביותר לדור הבא.
* שיטות שחלוף נמוכות, כמו ה- Single-Point, אשר יוצר ילדים הדומים במידה רבה להורים שלהם, או ערבוב שטוח של פרטים מוצלחים.
* בחירת הורים רק מהחצי הטוב של האוכלוסייה: פקטור בעל שני צדדים, למרות שהוזכר כמעודד Exploration, אולם הוא מכיל צד שמעודד גם Exploitation בכך שהוא מקדם החצי המוצלח מהאוכלוסייה על חשבון החצי האחר.

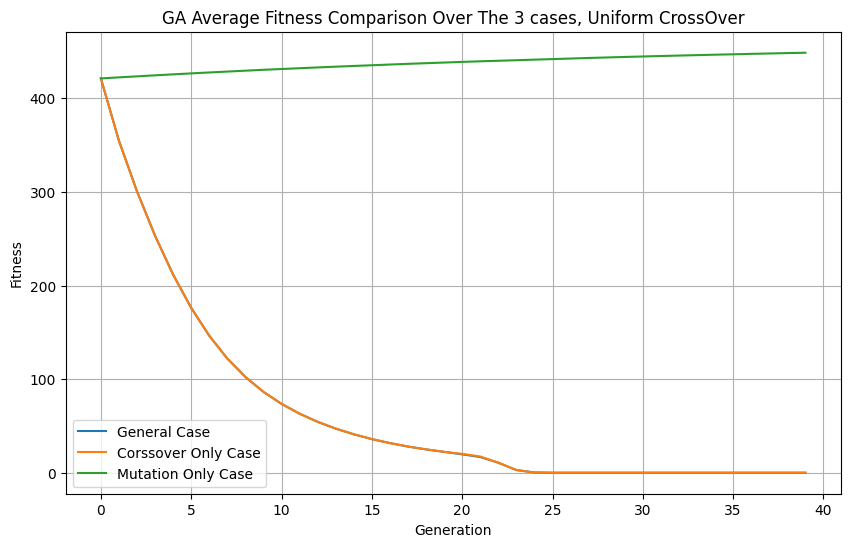
סעיף 6

ההשוואה בין 3 המקרים התבצעה לפי קריטריון הפיטניס הממוצע לאורך הדורות. הממוצע התבצע על פני 100 הרצות שבכל אחת מהם השתמשנו באותה אוכלוסייה התחלתית ב- 3 המקרים. האוכלוסייה ההתחלתית מורכת מפרטים בעלי גינומים שנוצרים באופן אקראי כפי שהדבר מתבצע באלגוריתם הרגיל.

כמו כן, נלקח בחשבון סוג ה- Crossover וההשוואה התבצעה לכל Crossover לחוד.



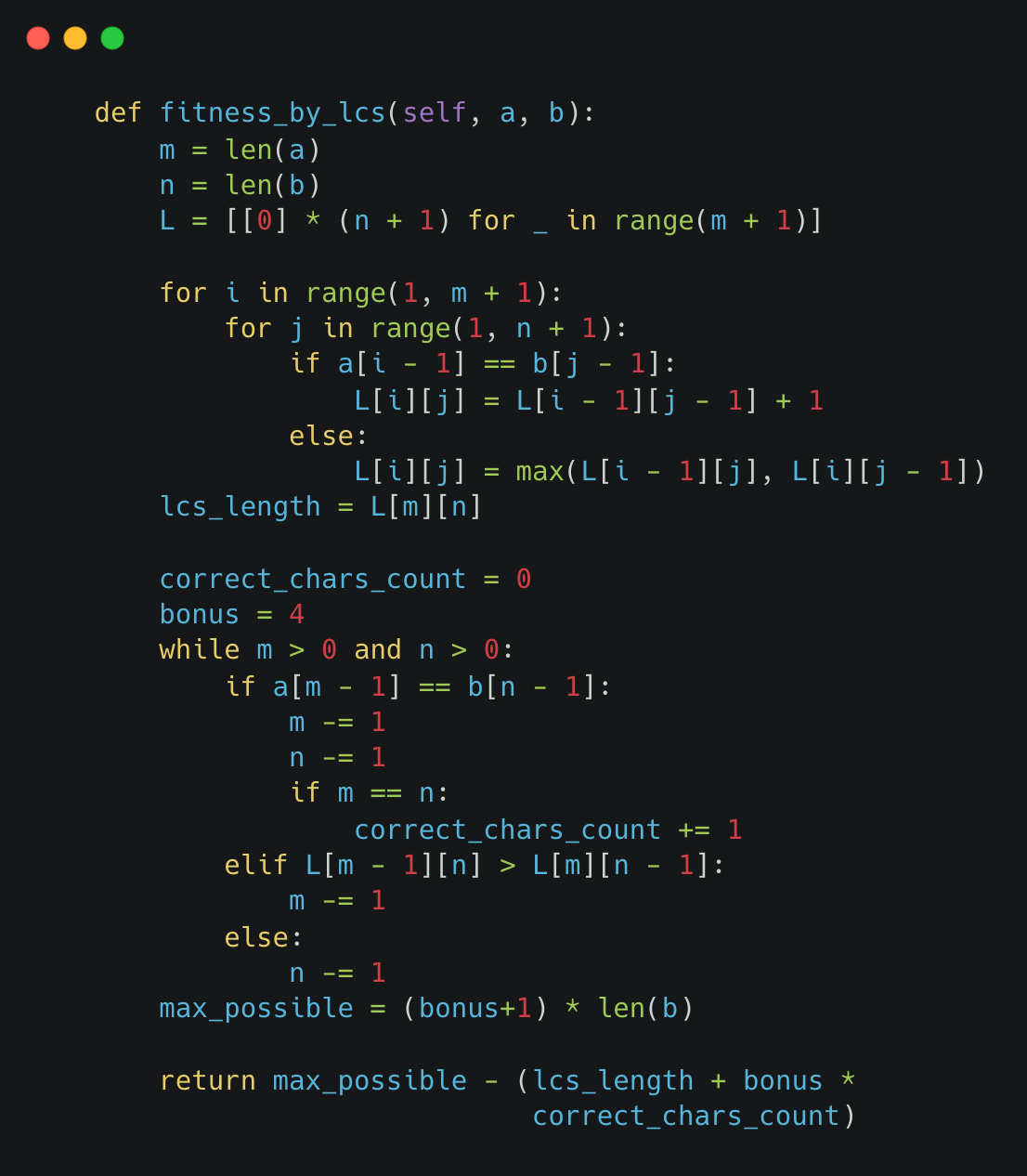




סעיף 7

התווספה האפשרות לבחור את הדרך בה הפיטניס מחושב מבין שתי השיטות: לפי המרחק ממחרוזת המטרה (DISTANCE), או לפי היורסטיקה המוצעת בסעיף (LCS), כאשר האחרונה מתבצעת ע"י הפונקציה המצורפת למטה, אשר משתמשת באלגוריתם תכנון דינמי על מנת למצוא את הרצף המשותף הארוך ביותר בין שתי המחרוזות. בנוסף, היא משתמשת ב- Backtracking על מנת לבדוק אם מיקומי האותיות ברצף הזה דומות בשתי המחרוזות, ומעניקה בונוס בהתאם.

הבונוס (2.0) נבחר על פני שני קריטריונים המצוינים לפי סדר העדיפויות: (1) המביא ליותר מקרים של התכנסות אופטימלית, (2) המתכנס יותר מהר (במונחים של מספר הדורות). הבונוסים השונים (0.0 עד 8.0) נבדקו על פני 100 אוכלוסיות שאותחלו באופן אקראי.



1. היוריסטיקה החדשה מקדמת פתרונות המכילים רצף נכון שהאותיות שלו במקומות הנכונים. בגלל שהאלגוריתם בוחר הורים מהפריטים הכי טובים, שהם לפי הגדרת היוריסטיקה הזו כאלה המכילים רצף אותיות הממוקמות נכון, קל יותר לאלגוריתם כעת ליצור משני הורים גינום שמכיל רצף יותר גדול ובעל מספר יותר גדול של אותיות הממוקמות נכון.
2. היוריסטיקה החדשה עדיפה על המקורית באופן מובהק. השיפור נצפה גם באחוז הפעמים בהם יש התכנסות לפתרון הנכון וגם בממוצע מספר הדורות הצריך עד להתכנסות. באשר לסיכוי ההתכנסות אין שיפור גדול מפני שתחת הפרמטרים האופטימליים האלגוריתם התכנס באחוז גבוה עוד לפני מימוש היוריסטיקה, אולם כן נצפה שיפור משמעותי במספר הדורות שעמד על ממוצע של 10.41 לעומת הטווח של 25-29 בהיוריסטיקה הקודמת.

סעיף 8

כדגכגד

סעיף 9