

מעבדה בבינה מלאכותית
203.3630 ב.1
סמסטר ב' – שנה"ל תשפ"ה

מרצה: שי בושינסקי

ניסוי מספר 1:

אלגוריתמים גנטיים (מערכות מתארגנות עצמית) - חלק א' במשימה

מועד הגשה אחרון: (יש להגיש יחדיו את שני החלקים של המשימה במועד זה)

יום ו' 10 באפריל 2025 – ההגשה באי-מייל ל- shay@cs.haifa.ac.il

מרכיב הציון:

התרגיל הינו חובה

תנאי ההגשה:

העבודה וההגשה בזוגות (ניתן כמובן להגיש ביחידים)

המשימה:

סקרנו אלגוריתם גנטי פשוט אשר חיפש מחרוזת נתונה תוך שימוש בהיוריסטיקה של מרחקי אותיות. המשימה שלכם במעבדה היא להרחיב ולשכלל את הדוגמא ולחקור את ביצועי האלגוריתם עפ"י השינויים הבאים:

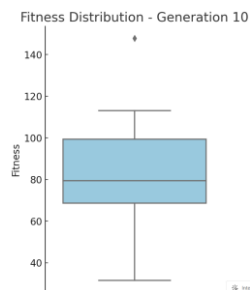
חלק א': יישום אלגוריתם גנטי ראשוני: הרחבת הדוגמא מההרצאה

1. הוסיפו חישוב ודווח בכל דור של ממוצע ה-FITNESS של האוכלוסייה ושל סטיית התקן מהממוצע, הפרט עם הפיטנס גרוע ביותר ואת טווח הפיטנס (הפרש בין המצטיין לגרוע)
2. הוסיפו חישוב ודווח בכל דור של זמן ריצה CLOCK TICKS וזמן ריצה אבסולוטי ELAPSED וכן עד להתכנסות למינימום לוקאלי או גלובאלי

3. הציגו שני גרפים של התנהגות הפיטנס מדור לדור עד להתכנסות:

a. השתנות הפיטנס מדור לדור (על אותו הגרף שלושה קווים בהתאמה לפיטנס של הפרט הטוב ביותר באוכלוסיה, הפרט הממוצע, והפרט הגרוע ביותר)

b. השתנות הפיטנס מדור לדור בBOXPLOT כלומר בכל דור תיבה המציינת את החציון בקו אופקי, גבולותיה מיצגות את הטווח הבינרבעוני, הקווים הקיצוניים את הטווח ובמעוינים ערכים חריגים. לדוגמא:



את הגרפים רצוי להפיק בתום הריצה של האלגוריתם הגנטי יש להסביר מה כל גרף מבטא בהקשר לריצה

4. ממשו כך שאפשר יהיה לבחור בין שלושת האופרטורים לשיחלוף SINGLE, TWO, UNIFORM

5. ציינו ונמקו אלו חלקים באלגוריתם אחראיים ל EXPLORATION ואילו ל EXPLOITATION?

6. השוו בין מצב בו אתם מריצים את האלגוריתם בקונפיגורציות הבאות

a. רק עם שיחלוף ללא מוטציות

b. רק עם מוטציות ללא שיחלוף

c. עם שניהם

סכמו מה המסקנות שלכם מריצות אלו

7. הוסיפו היוריסטיקה נוספת של LCS Longest Common Subsequence אורך הרצף המשותף הארוך ביותר עם מחרוזת היעד, כאשר הרצף שומר על סדר הופעת האותיות, ומעניקה בנוסף בונסים לאותיות הממוקמות במקומן הנכון.

השוו את ההיוריסטיקה הזו מול המקורית:

- a. הסבירו כיצד היא קשורה לשיחלופים במהלך האלגוריתם?
 - b. האם היא משפרת את ההיוריסטיקה המקורית ואם לאו מדוע?
- את ההשוואה יש לבצע תחת הפרמטריזציה המיטבית של האלגוריתם (גודל הבונס, גודל האוכלוסיה, קצב מוטציות וכו') – התיחסו לאיכות הפתרונות ולמהירות ההגעה אליהם

8. Exploitation Factor: הוסיפו שיטות למדידת לחץ הבחירה Selection Pressure ודווחו אותו מדור לדור

- a. Fitness Variance מתבצע על ידי מדידת הפיזור (השונות) של סיכויי הבחירה
 - b. Top-Average Selection Probability Ratio המתבצע על ידי השוואת ממוצע סיכויי הבחירה של תת-הקבוצה העליונה לסיכויי הבחירה של האוכלוסייה כולה.
- דווחו מדדים אלה בכל דור של אבולוציה וציינו מה כ"א מבטא

9. הוסיפו שיטות למדידת הגוון הגנטי Genetic Diversification (חלק זה מותאם בעיה) - Exploration Factor ודווחו אותם מדור לדור

- a. המרחקים בין הפרטים באוכלוסיה בדור הנוכחי
 - b. מספר האללים השונים באוכלוסיה בדור הנוכחי
 - c. האנטרופיה לפי שנון בדור הנוכחי
- דווחו מדדים אלה בכל דור של אבולוציה וציינו מה הם מבטאים ומה ההבדלים ביניהם. ממשו חישוב יעיל.

10. הוסיפו למנוע תמיכה בשיטות הבחירה שונות: parent selection

- a. RWS + Linear Scaling
- b. SUS + Linear Scaling
- c. טורניר דטרמיניסטי עם פרמטר K ו fitness ranking
- d. טורניר לא דטרמיניסטי עם פרמטרים K ו P ו fitness ranking
- e. הוסיפו שיטת שרידות נוספת Aging - לכל פרט באוכלוסיה יש מספר קבוע של דורות לשרוד
נתחו השפעות השיטות השונות כל הלחץ הגנטי והגיון

11. הוסיפו תמיכה למופע חדש של בעיה. בבעית הBIN PACKING: יש

לארוז עצמים בנפחים שונים במספר מיכלים בנפח V תוך שימוש במינימום מיכלים https://en.wikipedia.org/wiki/Bin_packing_problem

- a. גם לבעיה זו מצאו יצוג, פונקציית פיטנס ופונקציית פיטנס אדאפטיבית יעילים – נמקו בחירתכם

- b. הריצו את האלגוריתם שלכם על חמש בעיות בקובץ

binpack1.txt – הסבר על פורמט באתר

<http://people.brunel.ac.uk/~mastijb/jeb/orlib/binpackinfo.html>

והשוו את ביצועי האלגוריתם שלכם מבחינת מהירות ההתכנסות, איכות הפתרון וזמני ריצה לזה של אלגוריתם FIRST FIT החמדני

דוגמא לאלגוריתם FIRST FIT: $C=13$ $A=\{4,5,6,3,9,2\}$

$$B1 = \{4\}$$

$$B1 = \{4,5\}$$

$$B1 = \{4,5\} \quad B2 = \{6\}$$

$$B1 = \{4,5,3\} \quad B2 = \{6\}$$

$$B1 = \{4,5,3\} \quad B2 = \{6\} \quad B3 = \{9\}$$

$$B1 = \{4,5,3\} \quad B2 = \{6,2\} \quad B3 = \{9\}$$



ואת איכות הפתרון שהאלגוריתם מצאו והשוו מול המצוין בקובץ
כאופטימלי

12. הוסיפו תמיכה לבעיה נוספת : על המנוע שפיתחתם לפתור מבחר

חידות מתוך אתגר ה ARC של קאגל המתואר בקישור הבא:

<https://www.kaggle.com/c/abstraction-and-reasoning-challenge>

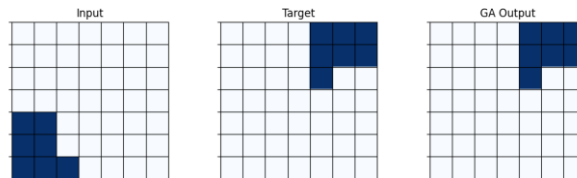
קבצי JSON עם אוסף החידות המלא תמצאו ב-

<https://github.com/fchollet/ARC-AGI>

חידות מתקדמות יותר ב-

<https://arcprize.org>

חידה פשוטה לדוגמא:



- עליכם למצוא יצוג מתאים לבעיה להגדיר פונקציית פיטנס מתאימה ולבחור אופרטורים גנטיים מתאימים.
- עליכם לתת למנוע לפתור את רשימת החידות הנתונה במובן של להתכנס לדפוס המוצג.

הנחיות:

- התרגיל ישמש כבסיס להרחבה ולתרגילים נוספים – לכן מומלץ להנדס אותו כך שחלקי הקוד יהיו שמישים שוב REUSABLE ושניתן לקנפג אותם לבעיות נוספות, ליצוג ידע שונה (מבנה נתונים) ולתתי אלגוריתמים אלטרנטיביים
- פרטו כיצד המדדים והגרפים שחושבו סייעו לכם בבחירת הפרמטריזציה המיטבית לפתרון הבעיות השונות בתרגיל

ההגשה:

במועד ההגשה יש להגיש דו"ח מסודר הכולל:

- א. תוכנת מקור SOURCE – מימוש הנ"ל בשפת תכנות לבחירתך (מתועדת) לפי הסטנדרטים של הקורס
- ב. תוכנות ריצה מתאימות EXE
- ג. מסמך המסכם את תוצאות הניסוי וניתוח רגישות לתוצאות – יש להתייחס לפרמטרים של האלגוריתם הגנטי, מס' הגנים הנדרשים, לשרידות של הגנים ולכל אספקט מעניין שתמצאו לנכון.