

תרגיל בית 2 – חישוב ביולוגי

מגשים:

בר אברהם דעבול ת.ז 324079243

נדב יוסף זדה ת.ז 208825257

בתרגיל זה כתבנו שתי תוכניות שקשורות למוטיפים בחישוב ביולוגי. יצרנו קובץ הסברים לשתי התוכניות והוא מצורף בהגשה כקובץ pdf בשם "תרגיל בית 2 - הסברים על התוכניות".

בנוסף, הכנו GitHub repository עם קובץ readme שמסביר איך להריץ את הקוד. הקישור לגיט הוא: https://github.com/BarDaabul/Biological_Computation_Project.git

שאלה 1

a. בשאלה זו נדרשנו לכתוב תוכנית שמקבלת כאינפוט מספר שלם חיובי n , ויוצרת את כל תתי-הגרפים הקשירים מגודל n . התוכנית תדפיס את הפלט בקובץ טקסט בפורמט המבוקש. כתבנו את התוכנית בשפת פייתון והיא נמצאת בקובץ "Ex2Q1.py".

b. בסעיף זה התבקשנו להריץ את התוכנית עבור ערכי n מ-1 עד 4. ביצענו את ההרצות הנ"ל. שמרנו את התוצאות בקבצי טקסט שצירפנו להגשה בתיקייה "שאלה 1-ב תוצאות מההרצות של התוכנית הראשונה". השמות של הקבצים הם בפורמט הבא: "subgraphs_n_i" כאשר i הוא הערך שנתנו ל- n באותה הרצה.

עבור $n=1$, קיבלנו תת-גרף אחד (צומת אחד עם לולאה עצמית).
עבור $n=2$, קיבלנו שני תתי-גרפים.
עבור $n=3$, קיבלנו 13 תתי-גרפים.
עבור $n=4$, קיבלנו 199 תתי-גרפים.

לכל אחת מההרצות האלה, קיבלנו תוצאה נכונה בהתאם למוטיפים הידועים עבור הערכים האלה.

סעיפים c ו-d:

ראינו כי עבור ערכי n מ-1 עד 4, אנחנו מצליחים להריץ את הקוד בזמן קצר מאוד. ניסינו להריץ את הקוד עבור $n=5$ במשך מספר שעות (כ-8 שעות), והקוד לא סיים את ההרצה. כלומר, עבור זמני הריצה המבוקשים בשאלה (שעה, שתיים, 4 שעות ו-8 שעות), התוכנית תסתיים בהצלחה עבור n שערכו המקסימלי הוא 4.

הסיבה לכך שהתוכנית לא מצליחה לרוץ בזמן סביר עבור n -ים גדולים יותר, היא שהבעיה של מציאת כל תתי-הגרפים היא בעיה קשה בעלת סיבוכיות זמן ריצה אקספוננציאלית. כתוצאה מכך, החל מערך מסוים של n הריצה של התוכנית לא יכולה להסתיים בזמן סביר.

שאלה 2

בשאלה זו נדרשנו לכתוב תוכנית שמקבלת כאינפוט מספר שלם חיובי n וגרף מכוון בפורמט נתון (גרף שמיוצג ע"י קשתות בקובץ טקסט). התוכנית יוצרת את כל תתי-הגרפים הקשירים מגודל n (כמו בשאלה 1), ובודקת לכל תת-גרף את מספר ההופעות שלו בגרף הנתון. התוכנית תדפיס את הפלט בקובץ טקסט בפורמט המבוקש.

כתבנו את התוכנית בשפת פייתון והיא נמצאת בקובץ "Ex2Q2.py".

לאחר שכתבנו את התוכנית הרצנו אותה על הגרף שניתן כדוגמה בשאלה (הגרף שמורכב מ-4 קודקודים, ומכיל את הקשתות $(1,2)$, $(2,3)$, $(1,4)$). הגרף ניתן לתוכנית כקלט ע"י קובץ טקסט. קובץ הטקסט שמייצג את הגרף שבחרנו כדוגמה נקרא "graph.txt" והוא מצורף בהגשה בתיקייה "שאלה 2 - קבצי טקסט".

הרצנו את התוכנית מספר פעמים, עם הגרף הנ"ל ועם ערכי n שבין 1 ל-4. התוצאות שהתקבלו מופיעים בקבצי טקסט ומצורפים להגשה בתיקייה "שאלה 2 - קבצי טקסט". השמות של הקבצים שמתקבלים מהתוכנית הם בפורמט הבא:

"SubgraphsWithCounting_n_i" כאשר i הוא הערך שנתנו ל- n באותה הרצה.

ראינו שעבור כל אחת מההרצות שביצענו, קיבלנו תוצאה נכונה בהתאם למוטיפים הידועים עבור ערכי ה- n ובהתאם למספר המופעים של כל אחד מהם בגרף שנתנו כקלט.