מבני נתונים פרויקט : ‏18/10/2020

בעבודה זו התבקשנו להשתמש במבני נתונים שונים על מנת לייצר קונקורדנציה שהיא אוסף של כלל המילים המופיעות בטקסט ומיקומן בו.

במהלך הכתיבה השתמשנו במספר מבני נתונים שונים אותם נפרט כעת בכדי למצוא את המתאים ביותר מבניהם למשימה הן מבחינת יעלות והן מהיבטי נוחות השימוש.

להלן המבנים השונים בהם השתמשנו:

1. רשימות מקושרות- השתמשנו במבני זה במחלקה lines מחלקה זו שומרת רשימה מקושרת של כלל השורות בהן מופיעה המילה (ללא כפילויות).
2. מחלקה DATA שומרת את המילה והרשימה המקושרת שמייצגת את מספרי השורות.
3. עץ חיפוש בינארי- השימוש במבני נתונים זה הינו בשמירת הdata- לפי סדר הABC .

הכנסת כל מילה לפי הסדר שלה בטקסט למקום הנכון במבני הנתונים לפי תכונותיו.

כל ערכי הצמתים בתת העץ הימני שלו גדולים מערך הצומת, וכל ערכי הצמתים בתת העץ השמאלי שלו קטנים מערך הצומת.

בחיפוש איבר בעץ חיפוש בינארי-

תחילה נבדוק את השורש ונשווה אותו לאיבר אותו אנו מחפשים, אם האיבר גדול יותר נעבור לבנו הימני, אם קטן יותר נעבור לבן השמאלי, אחרת-אם הוא שווה הרי שמצאנו את הערך המבוקש . אם סיימנו את העץ (הגענו לnull) הערך אינו נמצא בעץ

זמן הפעולה תלוי בגודל העץ (o)logn בממוצע ו (o)n בגרוע

1. הצגת המבנה הנתונים הינה לפי הפונקציה inorder אשר מציגה את הנתונים מהקטן לגדול החל מן התת עץ השמאלי הקיצוני והקטן ביותר דרך השורש של הצומת ועד לתת עץ הימני הקיצוני שהוא הגדול ביותר, כך אנו מקבלים את הסדר הרצוי לקונקורדנציה.

התכנית מקבלת קובץ ע"י fileChooser בחירת הקובץ ע"י המשתמש ושומרת את הקונקורדנציה בתוך תייקית הפרויקט- מומלץ לפתוח את הקובץ בפורמט של קבצי טקסט ולא בתוך סביבת הפיתוח.

\*הערה: אנחנו משתמשים בJAVAFX לכן יש לשנות את הJRE system library מjavaSE לJRE.

יעילות התוכנית:

יעילות הכנסה לעץ בינארי-

סיבוכיות ממוצעת להכנסה לעץ של איבר אחד הינו o(lgn)

אנחנו רוצים להכניס n איברים (מילים) לתוך העץ ולכן הסיבוכיות הינה **o(nlgn)**

הסיבוכיות של הכנסה לרשימה מקושרת הינה o(1) מכיוון שיש לנו מצביע לסוף הרשימה ולא מתבצע חיפוש ולכן לא משפיעה על הסיבוכיות.