**גיליון יבש מספר 3 - מבני נתונים 234218 חורף 2020/21  
  
מגישים:   
נועם יעקבי - 316600782   
עידו קולמן - 316465202   
  
שאלה 1**  
מבנה הנתונים שנשתמש בו מכיל טבלת ערבול דינאמית של הארגזים מכל המחסנים, מערך של המחסנים, כאשר כל תא במערך המחסנים מכיל עץ AVL ממוין של כל הנפחים השונים של הארגזים המתאימים למחסן זה, וכל צומת בעץ הנפחים מכיל ערימת מקסימום של הארגזים בעלי הנפח המתאים, את הארגז בעל המשקל הגדול ביותר בנפח המתאים לצומת זה ומצביעים לארגז הכבד ביותר בכל תת עץ שלו.

*- נאתחל מערך מחסנים באורך k, כפי שלמדנו ניתן לאתחל מערך עם ערך קבוע ב-, כאשר נאתחל כל תא במערך להכיל עץ AVL ריק – כפי שראינו אתחול עץ ריק נעשה גם כן ב-.   
נאתחל טבלת ערבול דינאמית מסוג chain hashing של ארגזים מכל המחסנים כאשר המזהה יהיה ה-id של כל ארגז, כפי שראינו אתחול של טבלת ערבול דינאמית ריקה נעשה גם כן ב-.  
לכן סיבוכיות הזמן הכוללת של הפעולה היא .*

*- נרצה להוסיף את הארגז למבני הנתונים המתאימים שלנו. לאחר בדיקת הקלט,   
ניצור משתנה מטיפוס ארגז המכיל את כל הנתונים שקיבלנו ובנוסף גם next ו-prev לטובת טבלת הערבול הדינאמית שלנו. נכניס את הארגז לטבלת הערבול הדינאמית שלנו מסוג chain hashing, ומכיוון שאנו דואגים לשנות את גודל הטבלה – פקטור העומס הוא ולכן כפי שראינו הכנסה נעשית בסיבוכיות בממוצע על הקלט, משוערך, ונשמור מצביע אל הארגז.  
כעת נרצה לשמור את הארגז במחסן המתאים לו. כל תא במערך מכיל עץ AVL של נפחים שונים במחסן, כאשר כפי שהזכרנו* כל צומת בעץ הנפחים מכיל ערימת מקסימום של הארגזים בעלי הנפח המתאים, ומצביע לארגז הכבד ביותר בעל נפח שווה או קטן לנפח המתאים לצומת. *עבור הכנסת הארגז החדש שלנו נחפש בעץ ה-AVL שלנו את המקום בו אמור להיות הנפח המתאים לו – כאשר ראינו כי חיפוש בעץ AVL ממוין בעל לכל היותר n נפחים שונים (מכיוון שיש לכל היותר n ארגזים במחסן) נעשה בסיבוכיות של , כאשר בכל צומת שנעבור בו נבצע בדיקה של מספר קבוע של פעולות – אם אנו פונים ימינה כלומר הנפח של הארגז החדש גדול מהצומת הנוכחי, נבדוק האם המשקל של הארגז החדש גדול מהמשקל של הארגז הקודם ששמור עבור תת העץ הימני בתור הכבד ביותר, במידה וכן נשמור את העץ החדש במקום, במידה ולא נמשיך בחיפוש. במידה ואנו פונים שמאלה, כלומר הנפח החדש קטן מהנפח של הצומת הנוכחי, נבדוק האם המשקל של הארגז החדש גדול מהמשקל של הארגז הקודם השמור בתור ההכי כבד בתת העץ השמאלי, במידה וכן נשמור בצומת זה את הארגז החדש בתור הכבד ביותר, ונבדוק האם הוא גם גדול יותר מהארגז בתת העץ הימני ונחליף במידה וכן. לאחר שמצאנו את המקום בו הנפח של הארגז אמור להיות – במידה והוא לא קיים ניצור צומת חדש ונאתחל אותו כנדרש, עם ערימת מקסימום שמכילה רק את המצביע לארגז החדש שלנו, ובמידה והעץ יצא מאיזון נעדכן את כל הצמתים המתאימים. במידה והנפח של הארגז כבר קיים בעץ, נכניס את הארגז לערימת המקסימום של הנפח – כפי שראינו נעשה בסיבוכיות של גובה העץ כלומר לכל היותר .   
לכן סיבוכיות הזמן הכוללת של הפעולה היא* *בממוצע על הקלט, משוערך.  
  
 – נרצה להעביר ממחסן i למחסן j את הארגז בעל המשקל הגדול ביותר בעל נפח שהוא לכל היותר V. ראשית נחפש במחסן ה-i בעץ ה-AVL שלו את הנפח V או את הנפח הכי גדול הקטן ממנו, ונבדוק במידע השמור בצומת מה הארגז הכבד ביותר בנפח זה ומה הארגז הכבד ביותר הנמצא בתת העץ השמאלי שלו – כלומר תת העץ בעל נפחים הקטנים ממנו ממש. במידה וקיים ארגז כבד יותר בתת העץ השמאלי שלו נחפש את הנפח המתאים לו עבור המידע הצמות בצמתים, אחרת נישאר בצומת הנוכחי. מכיוון שמידע זה שמור בצומת השוואה זו נעשית ב-. כמו כן נבצע בדיקה זו עבור כל הצמתים שעברנו בהם בחיפוש והנפח שלהם קטן או שווה לנפח V.  
לאחר שמצאנו את הצומת בו שמור הארגז הכבד ביותר בעל נפח שהוא לכול היותר V, נוציא את הארגז מערימת המקסימום של צומת זה – כפי שלמדנו הוצאת המקסימום מערימת מקסימום נעשית בסיבוכיות של .  
חיפוש הנפח המתאים ומציאת הנפח של הארגז הכבד ביותר נעשה בסיבוכיות של גובה העץ כלומר .   
לאחר שהסרנו את הארגז מערימת המקסימום של הנפח, נעדכן את המקסימום הנוכחי של הצומת, וכן נעדכן את כל הצמתים שמעליו שיבדקו אם הארגז הכבד ביותר השתנה מבחינתם. במידה וארגז זה היה אחרון בנפח המתאים לו, נסיר את הנפח מעץ ה-AVL של הנפחים, נבצע גלגולים במידה וצריך ונעדכן את הנתונים של מי ששינינו את מקומו או שהיה מעל לנפח שהסרנו כך שיצביע לארגז הכבד ביותר החדש. נעבור חזרה למעלה לפי אורך העץ על מנת לבצע תיקונים, כל תיקון נעשה ב-לכל גובה העץ ולכן פעולת התיקון תיקח גם כן .  
לאחר שהסרנו ארגז זה לחלוטין מכל הקשור למחסן i, נעדכן את המצביע כך שכעת כתוב שהוא שייך למחסן j,   
כלומר מכיוון שעבדנו עם מצביעים לא נצטרך לחפש את הארגז בטבלת הערבול הדינאמית של הארגזים אלא נוכל ישר לשנותו לאחר שמצאנו אותו, ונוסיף את הארגז החדש למחסן ה-j בדרך זהה לזו שתיארנו עבור פונקציית למעט ההכנסה לטבלת הערבול, הוספה שראינו כי נעשית בסיבוכיות של .  
לכן סיבוכיות הזמן הכוללת של הפעולה היא .  
  
  
 – אנו מחפשים את הקופסה בעלת המזהה id בטבלת הערבול שלנו, כפי שראינו חיפוש בטבלת ערבול דינאמית מסוג chain hashing, מכיוון שאנו דואגים להקטין ולהגדיל את הטבלה פקטור העומס הוא , ולכן חיפוש בטבלת הערבול נעשה בסיבוכיות של בממוצע על הקלט (ללא משוערך, מכיוון שלא נצטרך מעולם להגדיל או להקטין את הטבלה עבור חיפוש).  
מכיוון שבטבלת הערבול שמרנו את טיפוס הארגז יחד עם כל הפרטים המזהים שלו כולל המחסן אליו הוא היה שייך בהתחלה, ועדכנו את הארגז עבור כל העברה שלו ממחסן למחסן, נוכל ב-לגשת למידע של הארגז בו כתוב לאיזה מחסן הוא שייך.  
לכן סיבוכיות הזמן הכוללת של הפעולה היא* ***בממוצע על הקלט.***

**שאלה 2**  
**א.** מבנה הנתונים שנשתמש בו יהיה מבוסס על ערימות, *כאשר על מנת לממש את מבנה הנתונים ניצור רשימה מקושרת על מערכים בגודל k אשר עליהם יהיו מבוססים הערימות – מימוש ערימות כעץ כמעט שלם במערך, כאשר סדר הערימות ברשימה המקושרת יבטא את סדר ההגעה של האנשים כך שהאיבר הראשונה ברשימה המקושרת יהיה חדר ההמתנה ושאר האיברים – כלומר שאר הערמות ברשימה המקושרת, ידמו את התור.*

*- נאתחל עם הערך 1- מערך בגודל k עבור הערימה – אתחול אשר ראינו כי נעשה בסיבוכיות של .  
נאתחל את הרשימה המקושרת של התור כך שאיבר הראשון בה יהיה המערך שיצרנו, אתחול אשר נעשה ב-, ונשמור מצביע לתא הריק הבא, כאשר בהתחלה הוא יהיה התא הראשון במערך הראשון ברשימה.  
לכן סיבוכיות הזמן הכוללת של הפעולה היא .*

*- נוסיף אדם בגובה h לתא הריק אליו אנו מצביעים ונקדם את המצביע באחד במידה והמערך לא מלא, או במידה והמערך שהוספנו אליו התמלא ניצור מערך חדש, נוסיף אותו לרשימה המקושרת ונצביע על התא הראשון שלו, מספר פעולות קבוע ולכן נעשה בסיבוכיות של .  
במידה וכל המערך אליו הוספנו את האדם מלא, כלומר כל k התאים בו מלאים, ניצור ערימת מקסימום מהמערך המלא, פעולה אשר ראינו כי מתבצעת ב-.  
  
חישוב סיבוכיות משוערכת – עבור כל k קריאות לפעולה נבצע:*

* *k הכנסות למערך קיים וקידום המצביע לתא הריק הבא -*
* *יצירת ערימה מהמערך שהתמלא -*
* *יצירת מערך חדש במידה וצריך כאשר המערך הקיים התמלא ולא קיים איבר הבא ברשימה -*

לכן סיבוכיות הזמן המשוערכת לפעולה עבור k קריאות תהיה:

*כלומר קיבלנו כי סיבוכיות הזמן של הפעולה היא* ***משוערך.***

*- נקרא לאדם הגבוה ביותר מחדר ההמתנה, כלומר לאדם הנמצא בראש ערימת המקסימום הראשונה ברשימה המקושרת, ולאחריה נתקן את הערימה – תיקון שראינו כי נעשה ב-.  
במידה והערימה התרוקנה, כלומר הוצאנו כרגע את האדם האחרון מהערימה בחדר ההמתנה, נהפוך את הערימה או המערך הבאים ברשימה המקושרת להיות חדר ההמתנה החדש שלנו, כאשר נוכל לקרוא לאדם הבא בתור רק אם המערך כבר הפך לערימה. במידה ואין מערך או ערימה במקום הבא ברשימה המקושרת, ניצור מערך ריק חדש ונצביע עליו כעת. כל בדיקות ותיקונים אלו נעשים בסיבוכיות של .  
לכן סיבוכיות הזמן הכוללת של הפעולה תהיה .*

***ב.*** *נניח בשלילה כי קיים מימוש כזה.   
כלומר עבור מערך לא ממוין המכיל k אנשים, נבנה את מבנה נתונים מסעיף א' ב-, הוספנו את הגבהים של k האנשים ב-משוערך למערך – כלומר ב-k פעולות הכנסנו k גבהים של אנשים ולכן סיבוכיות כוללת של הכנסות היא .   
נניח כי הצלחנו לממש את פעולת ב-, כעת נוכל להוציא k גבהים ממוינים לפי סדר יורד, כלומר קיבלנו מערך ממוין בסיבוכיות של , זאת בסתירה לכך שלמדנו כי לא ניתן למיין מערך המכיל kבמיון מבוסס השוואות בסיבוכיות בזמן , ולכן הוכחנו כי לא קיים מימוש כזה.*

**שאלה 5**

*ניצור עץ סיומות מוכלל עבור n רצפי ה-DNA שקיבלנו, נרצה לשמור עבור כל צומת את מספר העלים בעלי ה- השונים בתת העץ שלו. לאחר מכן, עדיין בשלב ה-, נמצא את הצומת המכיל את הרצף ה-p מייצג עבור כל p אפשרי (שייתן לנו תוצאה שלמה עבור ) על מנת לייעל את פעולת , כאשר בפעולה יישאר לנו רק לעבור מהצומת עד לשורש ולשמור את המחרוזת מהקשתות.*

*- נאתחל את עץ הסיומות המוכלל של כל רצפי ה-DNA שקיבלנו בעזרת אלגוריתם הקופסה השחורה שנלמד בהרצאה כך שיצירת העץ נעשית ב- כאשר S הוא סכום אורכי רצפי ה-DNA.   
לאחר שבנינו את עץ הסיומות נבצע סיור postorder על כל העץ, עבור כל צומת פנימי נשמור מערך היסטוגרמות בגודל n, כאשר בכל תא נסמן כמה שונים קיימים בתת העץ שלו, ועבור עלה כמובן נאתחל למערך שכולו 0 למעט התא ה- iשבו נסמן אחד עבור ה- ששייך לקשת המובילה לעלה זה, עבור כל צומת פנימי נעדכן את מערך ההיסטוגרמות שלו על פי המערכים של הבנים שלו.   
תוך כדי עדכון המערך נשמור עבור כל צומת את מספר התאים השונים שיש לו מ-0, כלומר את מספר ה- השונים שנמצאים בעלים של תת העץ שלו. כל הסיור והפעולות בו יתבצעו ב-  
  
לאחר בניית העץ והמערכים בכל צומת, נשמור מערך חדש בגודל n+1 עבור מצביעים לצמתים בעץ.  
כל תא iבמערך יהיה מתאים לצומת בו נמצאת המחרוזת הכי ארוכה שהיא , כלומר הצומת אשר ממנו הקשת לשורש העץ תיתן לנו את רצף ה-DNA הארוך ביותר שהוא מייצג , אשר ייתן לנו את כל ה- האפשריים בהנחה ש-הוא בהכרח מספר שלם.  
עבור כל נמצא כעת מי הצומת שייתן לנו את המחרוזת הארוכה ביותר שהיא מייצגת.  
לשם כך נבצע סיור postorder בעץ הסיומות המוכלל שלנו, כעת נשמור מידע עבור אורך המחרוזת הנוכחית שאנו בודקים וכן עבור אורך המחרוזת המקסימלית שהיא מייצגת. עבור כל צומת שאנו עוברים בו נבדוק האם מספר ה- הוא גדול או שווה ל-(שמדמה כעת את ) בעזרת המידע ששמרנו בעדכון המערכים, במידה והוא גדול או שווה ל-נבדוק האם אורך המחרוזת שאנו בודקים כעת גדול מאורך המחרוזת המקסימלית ששמרנו כבר שהיא מייצגת, במידה וכן נשמור את המחרוזת הנוכחית, כלומר נשמור את המצביע לצומת המתאים וכן את אורך המחרוזת שלו.  
לאחר סיום סיור ה-postorder נשמור בתא ה-i את המצביע לצומת המכיל את רצף ה-DNA הארוך ביותר אשר הוא מייצג.  
מכיוון שביצענו מספר פקודות קבוע וסופי של פעולות השוואה בכל צומת סיבוכיות הזמן בכל צומת היא ,   
סיבוכיות הסיור מכיוון שעברנו על כל עץ הסיומות וביצענו פקודות עבור כל צומת תהיה .  
מכיוון שאנו מבצעים תהליך זה עבור כל , מילוי טבלת המצביעים לתאים תעלה לנו עבור כל i ולכן מכיוון שיש n+1 תאים שונים סיבוכיות המילוי תיקח .  
  
כלומר, בנינו עץ סיומות ב-, ביצענו סיור postorder ויצרנו מערכי היסטוגרמות עבור כל צומת בסיבוכיות של וכן בניית ומילוי מערך n המצביעים עבור רצפים מייצגים נעשה בסיבוכיות של   
לכן סיבוכיות הזמן הכוללת של הפעולה תהיה כנדרש*

*- נחשב את ,   
במערך המצביעים לצמתים שחישבנו בפעולת ה-ניגש לתא ה-, בהנחה ש-הוא מספר שלם אכן קיים תא כזה שכן המערך מכיל תאים מ-0 ועד n.  
נלך לצומת אליו שמור המצביע בתא ה-, על פי מה שחישבנו בפעולת ה- אנו יודעים כי המסלול מהצומת אל השורש ייתן לנו את הרצף הארוך ביותר שהוא p-מייצג. נשמור את המחרוזת השמורה בקשתות מהצומת אל השורש, כך שנקבל את המחרוזת הרצויה.  
ביצענו חישוב וגישה לתא במערך – נעשה בסיבוכיות של ,   
לאחר הגישה לצומת אליו מצביע התא במערך, עברנו על כל הקשתות והצמתים מהצומת אליו ניגשנו ועד לשורש ושמרנו את המחרוזת – מעבר הנעשה בסיבוכיות של לכל היותר אורך המחרוזת כלומר - .  
לכן קיבלנו כי סיבוכיות הזמן הכוללת של הפעולה תהיה .*