**שאלה 2**

מבנה הנתונים שלנו יכיל 2 מערכים שעובדים בשיטת שלמדנו בקורס, כך שעבור כל מערך נגדיר סדרת פונקציות ערבול שמתאימות לו. בכל עת, אחד המערכים ייקרא **המערך הראשי** והמערך השני ייקרא **המערך המשני** (יובהר טוב יותר בהמשך). סדרת פונקציות הערבול עבור כל מערך יוגדרו כפי שראינו בקורס:

* נגדיר כאשר הם גדלי המערכים בהתאמה.
* נגדיר עבור קטן כלשהו (לדוגמא ניקח (.
* לבסוף נגדיר את סדרת פונקציות הערבול להיות:

על מנת להשתמש בפונקציות הערבול בצורה פשוטה, נגדיר במבנה שלנו את הפעולה (הפונקציה):  
ובנוסף נגדיר במבנה את הפעולה (הפונקציה):

ככה שעבור כל וגודל מערך כלשהו נוכל לקרוא לפונקציות אלו ולקבל את תוצאת הערבול בקלות.  
2 פעולות אלו מתבצעות בסיבוכיות זמן (חישוב מתבצע ב-).

בנוסף, נצטרך לשמור משתנים ולציין נקודות חשובות נוספות:

* נגדיר 2 משתנים עבור כל מערך, שישמרו את כמות האיברים הנמצאים בכל מערך ואת הגודל הכולל של כל מערך, כל שנוכל לשלוף כל אחד מהם בסיבוכיות זמן ונעדכן אותם כאשר אנו מגדילים את גודל המערך ו\או מכניסים איברים למערך.
* בכל פעם שנכניס איבר למבנה הנתונים, ההכנסה תתבצע רק לאחד מן המערכים, ולכן נשמור דגל בשם שיגיד לנו לאיזה מערך נדרש לבצע את ההכנסה (המערך הראשי). עבור כל הכנסה של איבר חדש למבנה, אנו נכניס את האיבר החדש למערך הראשי, ובנוסף נעביר איבר מהמערך המשני למערך הראשי, כך שבסה"כ בכל הכנסה נכניס 2 איברים למערך הראשי. אם המערך הראשי יהיה מלא בזמן ההכנסה, אנו נגדיל את המערך המשני ונהפוך אותו להיות הראשי, כך שההכנסות הבאות יהיו אליו, ונעדכן את כך שיגיד לנו מיהו המערך הראשי בכל עת. יכול להכיל את הערך או את הערך , המייצגים מי הוא המערך הראשי הנוכחי.
* כפי שאמרנו קודם, בכל פעם שאנו מכניסים איבר חדש למערך הראשי, אנו מעבירים איבר נוסף מהמערך המשני אל המערך הראשי (כל פעם את האיבר הבא שנמצא בקצה המערך המשני). לכן, נשמור משתנה ששומר את האינדקס של האיבר הבא שיש להוציא מהמערך המשני, הנדרש להעברה למערך הראשי. בכל פעם שנעביר איבר מהמערך המשני אל המערך הראשי, נקטין את ב-1 כדי שיכיל את האינדקס של האיבר הבא שיש להעביר מהמערך המשני אל המערך הראשי.

כעת, נתאר את מהלך הפעולות על המבנה:

* Init(): נאתחל את מבנה הנתונים על ידי אתחול המערך לגודל קבוע (לדוגמא 100), ונעדכן שגודל המערך הוא 100 וכמות האיברים בו היא 0 בסיבוכיות (אתחול קבוע והשמה פשוטה למשתנים). נאתחל את להיות באורך , ונעדכן שגודל המערך הוא 200 וכמות האיברים בו היא 0 בסיבוכיות (אתחול קבוע והשמה פשוטה). נעדכן את להיות בסיבוכיות כדי לציין שהמערך הוא המערך שמבצעים אליו הכנסות כעת (המערך הראשי). לבסוף, נאתחל את ל--1 בסיבוכיות זמן , כדי לציין את העובדה שאין אף איבר ב- שצריך להעביר ל- עבור ההכנסות החדשות העומודת לבוא. סה"כ סיבוכיות הזמן והמקום הם .
* Insert(x): נבדוק את הערך הנמצא בדגל . נניח בה"כ שהערך הוא , ולכן ההכנסה ככל הנראה צריכה להתבצע למערך (אלא אם מלא ואז ההכנסה תהיה ל-). נבדוק מה הוא הגודל הכולל של וכמה איברים נמצאים כעת ב- בסיבוכיות זמן (אנו שומרים את ערכים אלו במשתנים). נחלק למקרים:
  + **אם אינו מלא (אם הכמות הנוכחית של האיברים בו קטנה מגודלו – נגלה זאת בסיבוכיות זמן כי אנו שומרים את גדלים אלו):**
    - **תחילה נציין כי ב- יש לפחות 2 מקומות פנויים. הסיבה לכך היא שאנו מתחילים כאשר הוא בגודל זוגי, ובכל פעם אנו מכניסים אליו 2 איברים. כאשר מלא אנו מגדירים את להיות פעמיים גודלו של , וכאשר מלא אנו מגדירים את להיות פעמיים הגודל של , ולכן גדל בכפולות של 2.**
    - נכניס את האיבר למערך על ידי שימוש בפונקציה

כפי שלמדנו בקורס, תוך קידום של עד אשר לא תהיה התנגשות איברים. בממוצע על הקלט, סיבוכיות הזמן להכנסה היא כפי שלמדנו.

* + - נגדיל ב-2 את המשתנה ששומר את כמות האיברים במערך כי הוספנו איבר חדש ל-. ההגדלה מתבצעת בסיבוכיות זמן .
    - אם מצביע על אינדקס חוקי, כלומר אם הוא גדול או שווה ל-0, נעביר את האיבר הנמצא באינדקס במערך אל המערך על ידי מחיקתו מ- והכנסתו ל- תוך שימוש בפונקציה

*ובמקרה של התנגשות נקדם את עד אשר לא תהיה התנגשות. נחסר 1 מ- כדי שיצביע על האינדקס של האיבר הבא שיש להעביר מ- ל-. בממוצע על הקלט, סיבוכיות הזמן למחיקה והכנסה היא כפי שלמדנו.*

*הכנסת איבר חדש ל- והעברת איבר מ- ל- מתבצעים בסיבוכיות זמן בממוצע על הקלט, כפי שהראינו, ובסיבוכיות מקום כי לא הקצינו זיכרון דינמי חדש.*

* + **אחרת, אם מלא, כלומר אם כמות האיברים ב- זהה לגודל המערך , זה אומר שמערך ריק – כלומר אין בו אף איבר ששייך למבנה (נוכיח בהמשך – (1)):**
    - נמחק את מערך בסיבוכיות זמן , כי הוא ריק.
    - נאתחל מחדש את מערך להיות בגודל בסיבוכיות זמן כאשר הוא גודל המערך (הסיבוכיות היא כאשר נשתמש בשיטה של אתחול מערכים ב- שלמדנו בקורס).
    - נעדכן את המשתנה השומר את גודל המערך להיות , ונעדכן את המשתנה ששומר את כמות האיברים ב- להיות 2 כי אנו עומדים להכניס אליו 2 איברים חדשים. סיבוכיות הזמן לביצוע פעולות אלו הוא .
    - נכניס את האיבר למערך על ידי שימוש בפונקציה

כפי שלמדנו בקורס. לא יכולה להיות התנגשות כי מערך הוא מערך שאין בו אף איבר, ולכן נבחר ב-. סיבוכיות זמן ההכנסה אם כן היא .

* + - נעביר את הגודל למשתנה כדי לציין את העובדה שבהכנסה הבאה האיבר שאותו יהיה להעביר מ- ל- נמצא באינדקס במערך .
    - נמחק את האיבר הנמצא באינדקס במערך , ונעביר אותו למערך על ידי שימוש בפונקציה

כפי שלמדנו בקורס. אם יש התנגשות עבור , נשתמש ב-. לכן סיבוכיות הזמן של ההכנסה היא .

* + - נחסר 1 מ-, כך שיהיה שווה ל- כדי לציין את העובדה שבהכנסה הבאה האיבר שאותו יהיה להעביר מ- ל- נמצא באינדקס במערך , ונקטין ב-1 את המשתנה ששומר את כמות האיברים הנמצאים ב- מכיוון שהסרנו איבר 1 מ-.

יצירת המערך , הכנסת האיבר אליו והעברת איבר מ- ל- מתבצעים בסיבוכיות זמן . אם היה איברים במבנה לפני יצירת , סיבוכיות המקום היא , ולאחר יצירת סיבוכיות המקום תישאר כי גודלו של הוא והכנסנו בסה"כ איבר אחד נוסף למבנה.

* + - נוכיח את (1): לפי האלגוריתם שהצענו, נניח בה"כ ש- מלא. אם מלא וגודלו הוא , זה אומר שגודל המערך הוא , כי בכל פעם שמערך אחד מתמלא אנו מגדירים את גודלו של השני להיות פעמיים הגודל של המערך המלא. לכן, אם מלא, זה אומר שיש בו איברים, כי על מנת שהוא יתמלא אנו צריכים לבצע פעולות הכנסה, כאשר כל אחת מהן מכניסה איבר חדש ל- ומעבירה איבר נוסף מ- ל-. לאחר פעולות הכנסה, אנו בעצם מוציאים איברים מ-, ולכן כאשר מלא ריק בוודאות.
* : נשלוף את גודלם של ו-, ועבור כל אחד מהם נחשב את פונקציית הערבול המתאימה:

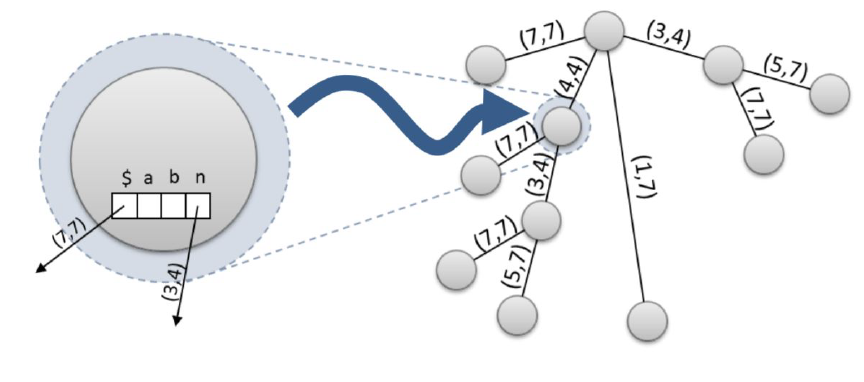
*נחפש את בכל אחד מהמערכים בסיבוכיות זמן בממוצע על הקלט, כפי שלמדנו בקורס. אם מצאנו את ב-1 מהם, נחזיר אותו.*

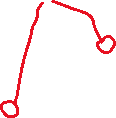
**נוכיח את סיבוכיות המקום:**

נניח ש-2 המערכים במבנה הם מאורך ו-. אנו מבצעים הגדלה למערך מסוים רק אם המערך השני התמלא. לכך המערך באורך קיבל את אורכו לאחר מילוי המערך באורך באיברים. מכיוון שאנו לא מבצעים הסרה של איברים, בכל עת במבנה יש לנו לפחות איברים (כי אחרת לא היינו מבצעים הגדלה למערך השני להיות באורך ). בנוסף, יש לנו לכל היותר איברים במבנה כי אחרת היינו מבצעים הגדלה של המערך שאורכו להיות באורך , וזו סתירה לכך שהגודל שלו הוא . כלומר יש לנו לפחות ולכל היותר איברים במבנה. בנוסף, קיימים לנו כמות משתנים סופית ששומרת פרמטרים נוספים, ולכן עבור איברים במבנה סיבוכיות המקום היא כי אנו שומרים מערכים שאוכם הכולל הוא לכל היותר ועוד כמות סופית של משתנים לא דינמיים.

**שאלה 4**

יהי עץ סיומות דחוס המייצג את הסיומות של מחרוזת כלשהי. אזי, כל קשת היוצאת משורש העץ מייצגת התחלה של סיומת בעץ, כך שכל הקשתות מייצגות את התחלתן של כל הסיומות הקיימות במחרוזת. בנוסף, כל עלה בעץ מייצג סיומת אחת בלבד, כפי שלמדנו. הקשת שבה האינדקסים הם מייצגת את תת המחרוזת במחרוזת , המתחילה בתו שהאינדקס שלו הוא ומסתיימת בתו שהאינדקס שלו הוא במחרוזת . לכן כדי לקבוע את אורך של סיומת כלשהי נצטרך לסכום את אורכי כל תתי המחרוזות המוכלות בסיומת ומיוצגים על ידי הקשתות הנמצאות במסלול של הסיומת. עבור כל צומת שהאינדקסים שלה הם , ההפרש מייצג את כמות התווים של התת מחרוזת אותה הקשת מייצגת. אם נקבל שהאורך הוא 1, כלומר הקשת מייצגת תו בודד. אם נקבל ש- כלומר הקשת מייצגת 2 תווים, וכו'. לכן, נגדיר את **אורך המסלול של סיומת** כסכום ההפרשים עבור כל קשת שהאינדקסים בה הם , אשר נמצאת במסלול היוצא מהשורש ומסתיים בעלה המייצג את הסיומת. לדוגמא:

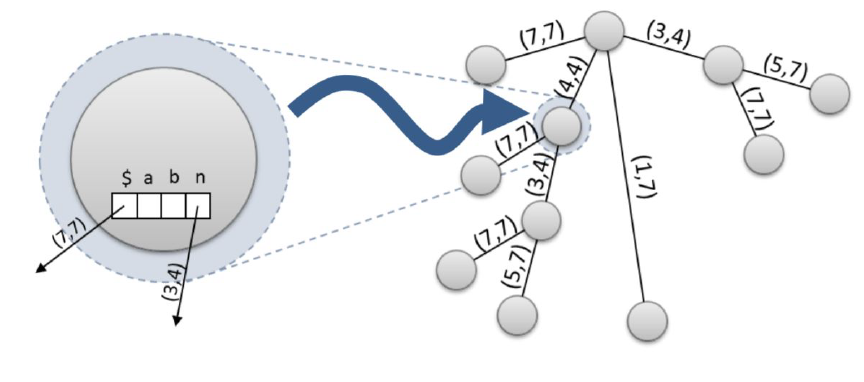




* אורך מסלול הסיומת של הסיומת המיוצגת על ידי העלה האדום הוא (4-4+1)+(4-3+1)+(7-5+1)=6. במקרה של מספר זה מייצג את אורך הסיומת , שאורכה הוא 6.
* אורך מסלול הסיומת של הסיומת המיוצגת על ידי העלה הכחול הוא (4-3+1)+(7-7+1)=3. במקרה של מספר זה מייצג את אורך הסיומת שהוא כמובן 3.

**כעת נתאר את מהלך האלגוריתם:**

* נבנה מערך של תווים באורך בשם , אשר בסוף יכיל את המחרוזת אותה אנו משחזרים. סיבוכיות הזמן לאתחול היא .
* לכל קשת היוצאת מהשורש אל בן שנסמנו , המסמלת את תחילתה של סיומת המתחילה בתו שנסמנו :
  + נבנה מערך של מספרים באורך בסיבוכיות זמן .
  + נבנה מצביע למספר שמייצג את כמות המספרים הנוכחית במערך, ונאתחל את ל-0 כדי לציין את העובדה שכעת אין אף איבר במערך .
  + **נכניס ל- את כל אורכי מסלולי הסיומות היוצאים מהשורש, עוברים דרך , ומגיעים לעלה (נסביר בהמשך כיצד נבצע זאת – (1), ונראה שסיבוכיות הזמן לבצע זאת היא ).** ייתכנו בסה"כ עלים מכיוון שיש בסה"כ סיומות, ולכן יהיו לכל היותר מסלולים היוצאים מהשורש, העוברים דרך , ומגיעים לעלה. כלומר, המערך לכל היותר יתמלא, אך לא תיתכן שגיאה הקשורה לגודלו. המערך , המכיל כעת את כל אורכי מסלולי הסיומות, בעצם מכיל את כל אורכי הסיומות המתחילות בתו . לדוגמא, עבור הקשת המסומנת באדום:





המערך יכיל את המספרים , כאשר 2 מייצג את המסלול האדום ואז הסגול (אורך הסיומת המתחילה בתו ), 6 מייצג את המסלול האדום ואז התכלת (אורך הסיומת המתחילה גם בתו ), ו-4 מייצג את המסלול האדום ואז הכתום (אורך הסיומת המתחילה גם בתו ). אורכים אלו, כפי שציינתי קודם, מייצגים את כל האורכים של כל הסיומות המתחילות בתו במחרוזת .

אורכי המסלולים מייצגים את אורכי הסיומות, ולכן נמקם את התו באינדקסים הרלטיביים לסוף המערך . כלומר, לכל מספר הנמצא במערך , נמקם את התו במקום במערך . למשל עבור הדוגמא לעיל, עבור , ולכן נמקם את התו באינדקסים שהם במערך . נקבל כעת ש- הוא:

סימני השאלה מייצגים את התווים שעוד לא גילינו, אשר נגלה כאשר נעבור על הקשתות הנוספות היוצאות מהשורש.

* + מכיוון שכל סיומת אפשרית במחרוזת מתחילה בקשת היוצאת מהשורש, אחרי שנעבור על כל הקשתות היוצאות מהשורש נקבל את המילה המשוחזרת ב-, מכיוון שמילאנו את מערך זה בכל התחלות הסיומות של המחרוזת אותה אנו משחזרים – כלומר בכל התווים הנמצאים במחרוזת.

1. **נסביר כיצד נכניס ל- את כל אורכי מסלולי הסיומות היוצאים מהשורש, עוברים דרך , ומגיעים לעלה:** נעשה מעבר רקורסיבי על כל הבנים של , כך שנעביר ברקורסיה את המערך , את המצביע שמצביע למספר המייצג כמה איברים נמצאים במערך , ואת אורך מסלול הסיומות הנוכחי. בכל בן ברקורסיה נוסיף לאורך מסלול הסיומות את התרומה של הבן לאורך המסלול, וכאשר נגיע לעלה (תנאי העצירה שלנו) נוסיף את תרומתו לאורך המסלול, נכניס את הסכום למקום הבא במערך , ונקדם את הערך שאליו מצביע ב-1, כדי לציין שגודל המערך גדל ב-1 ועל מנת שהעלה הבא יכניס את אורך המסלול אליו הוא שייך לאינדקס הבא במערך . כשנסיים את תהליך ה- ונחזור מהרקורסיה , נוסיף את אורך הסיומת המיוצגת על ידי הקשת היוצאת מהשורש אל לסכום עבור כל איברי המערך (כלומר נוסיף את תרומתה של הקשת היוצאת מהשורש אל אל כל אורכי הסיומות היוצאים מקשת זו) בסיבוכיות זמן (סריקה לינארית ועדכון הערכים) ונחזור אל השורש. הוכחנו בקורס שכמות הצמתים בעץ סיומות של מחרוזת היא , ולכן בסה"כ נקבל שסיבוכיות הזמן לסריקה היוצאת מכל בן של השורש היא .

לשורש יש מספר בנים שהוא לכל היותר מספר קבוע - , ולכן לבצע את כל סריקות ה- עבור כל אחד מהבנים מתבצע בסיבוכיות זמן כי כל סריקת בעצמה מתבצעת בסיבוכיות זמן , והרי שיש לנו מספר קבוע של סריקות כאלו.

**לסיכום, במחרוזת תימצא המחרוזת המשוחזרת, וביצענו את האלגוריתם בסיבוכיות זמן . בסה"כ הקצינו מערך חדש באורך ועומק הרקורסיה הוא לכל היותר במהלך ריצת האלגוריתם, כי אורך הסיומת הארוכה ביותר היא אורך המחרוזת עצמה, . לכן סיבוכיות המקום גם תהיה .**