

מעבדה 9. נושא: טבלאות גיבוב

תאריך הגשה: 10/01/2023 בשעה 23:00 (בזוגות)

יש לקרוא היטב לפני תחילת העבודה !

מבוא:

במעבדה הנוכחית נממש טבלת גיבוב.

תיאור:

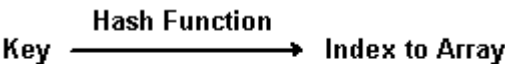
טבלת גיבוב היא מבנה נתונים שנותן גישה לרשומה באמצעות המפתח המתאים לה. המבנה עובד באמצעות הפיכת המפתח על ידי פונקציית גיבוב, למספר המייצג אינדקס במערך שמפנה אל הרשומה המבוקשת.

למשל: במפעל של 100 איש הוחלט לתת לכל עובד מספר ייחודי בן 3 ספרות. באופן כזה ניתן לשמור את כל העובדים במערך של 1000 איברים.
אולם, אם משתמשים בת.ז. של העובדים בני 9 ספרות, נצטרך מערך בגודל מיליארד (10^9) כדי להחזיק 100 עובדים בסה"כ.
ולכן רצוי להשתמש במבנה נתונים שונה ממערך בו עדיין אפשר לבצע חיפוש מהיר.

יתרונות:

חיפוש מהיר: בהינתן מפתח נתון (למשל שם של אדם), מצא את הרשומה המתאימה (למשל מספר הטלפון של אותו אדם).

פונקציית גיבוב:



יש הרבה סוגים של פונקציות גיבוב.
דוגמא 1 לפונקציית גיבוב כאשר המפתח הוא מספר שלם:
נתונים מספרים שלמים. יש למפות אותם למערך בגודל N
 $F(\text{key}) = \text{key} \% N$
למשל, עבור מספרים: 123450, 123467, 123456
ומערך בגודל N=10
נבצע:
 $123456 \% 10 = 6$
 $123467 \% 10 = 7$
 $123450 \% 10 = 0$

0	123450
1	
2	
3	
4	
5	
6	123456
7	123467
8	
9	

דוגמא 2 לפונקציית גיבוב כאשר המפתח הוא אינו מספר שלם:



למשל, אם נתונים השמות:

Sarah Jones
Tony Balognie
Tom Katz

בשלב ראשון, נעביר את השמות למספרים שלמים ע"י פונקציית מעבר ממחרזות למספר שלם. דוגמא לפונקציה כזאת יכולה להיות, אם נצמיד לכל תו את קוד ה-ASCII שלו ונחבר את כל הערכים יחד. התוצאה תהיה:

Sarah Jones --> 1038
Tony Balognie --> 1259
Tom Katz --> 746

בשלב שני, נשתמש בפונקציית הגיבוב מהדוגמא הקודמת :

Sarah Jones --> $1038 \% 10 \rightarrow 8$
Tony Balognie --> $1259 \% 10 \rightarrow 9$
Tom Katz --> $746 \% 10 \rightarrow 6$

התוצאה תהיה:

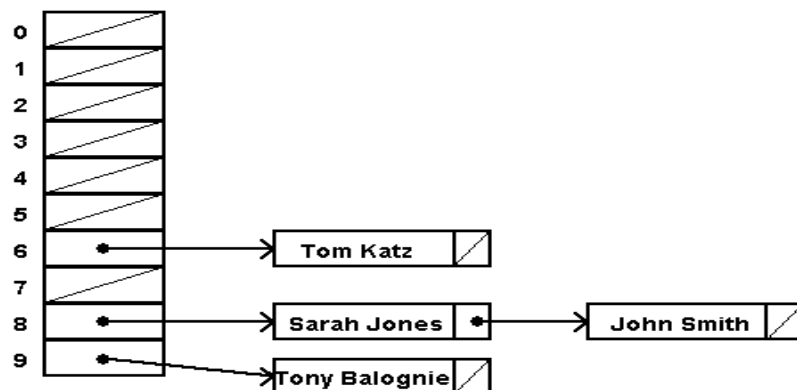
0	
1	
2	
3	
4	
5	
6	Tom Katz
7	
8	Sarah Jones
9	Tony Balognie

טיפול בהתנגשויות:

לעתים, יתכן כי שני מפתחות שונים יופנו לאותו התא במערך.

למשל, אם נוסף לדוגמא 2 : $8 : 948 \% 10 \rightarrow$ John Smith . נקבל כי תא מספר 8 כבר תפוס.

במעבדה הנוכחית, נטפל בגישת השרשרות (Chaining) לטיפול בהתנגשויות לפיה, במקרה של התנגשות, המפתחות ישורשרו ברשימה מקושרת. למשל, עבור דוגמא 2 שלעיל,



מישימה:

צרו מחלקה המגדירה **טבלת גיבוב** (HashTable) המכילה את השיטות contains, add, remove, clear, isEmpty. תיעוד המחלקה נתון לכם.

מחלקה זו דומה במעט למחלקה HashSet המממשת את Set:

<https://docs.oracle.com/javase/8/docs/api/java/util/Set.html>

הדרך התקנית להפוך מפתח כלשהו למספר שלם הוא להשתמש בשיטה hashCode. ניתן ליצור פונקציה כזאת באופן אוטומטי ב ECLIPSE.

הדרך התקנית לבדוק שוויון של עצמים היא להשתמש בשיטה equals. ניתן ליצור פונקציה כזאת באופן אוטומטי ב ECLIPSE.

ניתן להשתמש בפתרון של מעבדות קודמות, אבל לא בספרייה java.util.

המשימה (איך לא?) היא לגרום למחלקת הבדיקה HashtableTest לעבוד בצורה תקינה.

1. הוסיפו למחלקה Person מתודה equals המשווה בין שלושת השדות של Person. הוסיפו מתודת hashCode מתאימה. שימו לב שהשוואה זו שונה ממה שהיה במעבדה של עצי חיפוש. המחלקה Person נמצאת בחבילה: `package il.ac.telhai.ds.misc;`
2. ממשו את HashTable הנתונה, תוך שימוש במחלקה DLinkedList שמימשנו במעבדות קודמות (היא נתונה לכם עם מימוש המתודות הנדרשות). המחלקה DLinkedList והממשק List נתונים בחבילה `il.ac.telhai.ds.linkedlist`. המחלקה HashTable בחבילה `il.ac.telhai.ds.hash`. יתכן שבבנאי תקבלו "הערת אזהרה". כדי שהקומפיילר לא יתייחס לאזהרה זו, ניתן לכתוב מעליה:
`@SuppressWarnings("unchecked")`
אם עדיין יש אזהרה בבודק האוטומטי - ניתן להחליף שורה זו בשורה הבאה:
`@SuppressWarnings({"unchecked","rawtypes"})`

סדר העבודה ופרטים טכניים

- שליפת הפרויקט DS-Lab09-HashTable מתוך GITHUB בקישור:
<https://github.com/ykanizo/DSLAb2022-2023Public>
 - אם אין לכם גישה לפרויקט שהורדתם מ GITHUB במעבדות הקודמות יש לבצע שליפה מחדש.
 - אם יש לכם גישה לפרויקט שהורדתם מ GITHUB במעבדה הראשונה אז בצעו:
 - קליק על שם הפרויקט.
 - עכבר ימני
 - Team-->Pull

File-->Import->Git->Projects From Git->Existing Local Repository ■

פורמט קובץ ההגשה ובדיקתו:

פורמט: יש להגיש קובץ ZIP בשם

43_lab09_123456789_987654321.zip

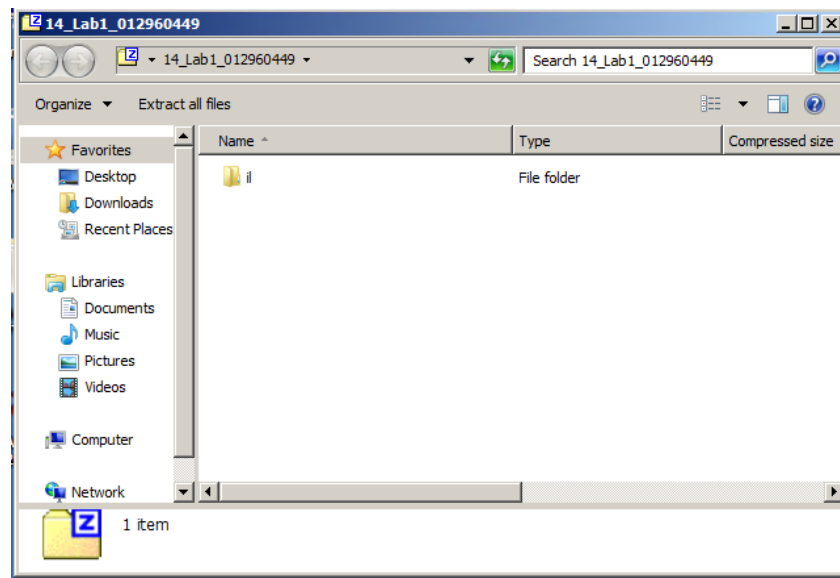
(כמובן, יש להחליף את המספרים עם מספרי ת.ז. של המגישים).

על הקובץ להכיל את כל קבצי ה JAVA שכתבתם כאשר הם נמצאים בתתי תיקיות בתוך התיקייה
il/ac/telhai/ds/

על פי המבנה של הפרוייקט הנתון.

כלומר, השורש של קובץ ההגשה יכול רק תיקייה בשם il, והוא יכול את כל קבצי - java על פי
התבנית הנתונה בפרוייקט.

להמחשה תמונה של קובץ כזה שנפתח ב - WindowsExplorer



בדיקת קובץ ההגשה: בדקו את הקובץ שיצרתם בתוכנת הבדיקה בקישור:

<https://csweb.telhai.ac.il/>

ראו [סרטון הדגמה](#) של השימוש בתוכנת הבדיקה.

חשוב !!!

בדיקת ההגשות תבוצע ברובה ע"י תוכנית הבדיקה האוטומטית הנ"ל. תוצאת הבדיקה תהייה בעיקרון זהה לתוצאת הבדיקה הנ"ל שאתם אמורים לערוך בעצמכם. כלומר, אם ביצעתם את הבדיקה באתר החוג, לא תקבלו הפתעות בדיעבד. אחרת, ייתכן שתרגיל שעבדתם עליו קשה ייפסל בגלל פורמט הגשה שגוי וכו'. דבר שהיה ניתן לתקנו בקלות אם הייתם מבצעים את הבדיקה. היות ואין הפתעות בדיעבד, לא תינתן אפשרות של תיקונים, הגשות חוזרות וכד'.

הגשה שלא מגיעה לשלב הקומפילציה תקבל ציון 0.

הגשה שלא מתקמפלת תקבל ציון נמוך מ- 40 לפי סוג הבעיה.

הגשה שמתקמפלת תקבל ציון 40 ומעלה בהתאם לתוצאות הריצה, ותוצאת הבדיקה הידנית של הקוד (חוץ ממקרה של העתקה).

תכנית הבדיקה האוטומטית מכילה תוכנה חכמה המגלה העתקות. מקרים של העתקות יטופלו
בחומרה