

## תרגיל בנושא Hashing (נתנאל)

## הוראות הגשה

שאלות בנוגע לתרגיל נא להפנות דרך ה-google-group בה השתמשתם בקורס מבוא למדעי המחשב:

<http://groups.google.com/group/89-110>

אם לא נענתה תשובה תוך 24 שעות (לא כולל שבת), נא לשלוח אלי (נתנאל) מייל עם לינק לדיון הרלוונטי ואענה. בכל מייל יש לציין שם, שם משתמש, מס' קורס, וקבוצת תרגול.

**הנהלים זהים לאלה של הקורס 89-110/88-170.**

**יש לציין בתחילת הדיון: 89-120.**

- מועד פרסום: 29/05/14
- מועד אחרון להגשה: 23:00 12/06/14
- ערך התרגיל: 25% מציון התרגיל
- יש לשלוח את הקבצים באמצעות האתר: <http://help.cs.biu.ac.il/submit.htm>
- לפני חלוף התאריך הנקוב לעיל.
- שם ההגשה של התרגיל: HashingEx
- יש להקפיד מאוד על כל הוראות עיצוב הקלט והפלט, כמפורט בכל סעיף וסעיף. על הפלט להיראות בדיוק כמו בדוגמאות. אין להוסיף או להשמיט רווחים או תווים אחרים ואין להחליף אותיות גדולות בקטנות או להיפך ⊗ אי-הקפדה על פרטים אלה עלול לגרור ירידה משמעותית ביותר בציון התרגיל עד כדי 0. ראו הוזהרתם!
- להזכירכם, העבודה היא אישית. "עבודה משותפת" דינה כהעתקה.
- אין להדפיס למסך שום דבר מעבר למה שנתבקש בתרגיל.
- יש לוודא שהתרגיל מתקמפל ורץ על השרתים באוניברסיטה (u2) ללא שגיאות/אזהרות.

**על פי הוראת המחלקה למדעי המחשב סטודנטים העושים את הקורס 89-111 צריכים לעשות את התרגיל בשפת C.**  
**לסטודנטים בקורס 88-174, מומלץ לעשות את התרגיל ב++C.**

## רקע

לצורך תכנות התרגיל, חשוב מאוד שתעברו על המצגת הרלוונטית מהאתר שלי:  
[/http://u.cs.biu.ac.il/~gelemn/89120](http://u.cs.biu.ac.il/~gelemn/89120)

**שאלות שהתשובה אליהן מופיעה במצגת לא ייענו! עברו עליה היטב.**

## **שימו לב!**

**עליכם לבדוק היטב את התרגיל בסביבת ה-u2.  
בהוראת מרצה הקורס תהיה רק בדיקה אחת ולא יינתנו תיקונים, ולא משנה האם התרגיל עבד לכם בסביבות אחרות.**

## Hashing (in C/C++)

### הנחיות עבור התרגיל

- שם התרגיל: HashingEx
- עבור המגישים בשפת C: יש להגיש קובץ makefile וכן שלושה קבצים: funcs.c, main.c ו-1- funcs.h. התוכנית שלכם תקומפל על ידי ה-makefile שלכם.
- עבור המגישים בשפת C++: יש להגיש קובץ makefile, קובץ main.cpp וכן קבצי H-1 CPP לכול מחלקה. התוכנית שלכם תקומפל על ידי ה-makefile שלכם.

### הבהרה: לצורך התרגיל מותר לכם להשתמש בספריות הנדרשות לצורך הקצאות דינאמיות.

לצורך בחירת מצבה וערך שיסמן תא ריק, אתם יכולים להשתמש במערך ששמור את התאים הריקים, או להגריל מספר רנדומי גדול וזה יעבוד לכם בהסתברות של יותר מ- $1 - (1/1000000)$ , במידה ולא יוכנסו הערכים שבחרתם כקלט (אני הייתי לוקח את הסיכון) והתוכנית תעבוד כמו שצריך.

### זרימת התוכנית:

1. קלוט מספר שלם כסוג ה-hash. ניתן להניח שיתקבל אחד משלושת הערכים הבאים: 0,1,2.
    - a. במידה ונקלט 0: מדובר ב-Chaining.
    - b. במידה ונקלט 1: מדובר ב-Linear Probing עם הסרה לפי שיטת "צמצום רווחים".
    - c. במידה ונקלט 2: מדובר ב-Double-Hashing עם הסרה לפי שיטת המצבה.
  2. קלוט את גודל טבלת ה-hash (ניתן להניח שיתקבל ערך שלם גדול מ-3).
  3. קלוט פעולות (0,1,2,3,4). ניתן להניח שיתקבלו אחד מהערכים הללו כמספר פעולה) ובצע אותן על פי התפריט הבא, כל עוד לא נקלטה פעולה המציינת את סיום התוכנית.
    - a. במידה ונקלט 0: סיים את התוכנית.
    - b. במידה ונקלט 1: קלוט מספר שלם נוסף, והוסף אותו לטבלת ה-hash.
    - c. במידה ונקלט 2: קלוט מספר שלם נוסף, והסר אותו מטבלת ה-hash במידה והוא נמצא בה. אם לא – לא עושים כלום.
    - d. במידה ונקלט 3: קלוט מספר שלם נוסף, וחפש אותו בטבלת ה-hash. יש להדפיס את כל הערכים בהם נתקלים בדרך (כולל מצבה במקרה הרלוונטי, על ידי X), כולל המספר עצמו במידה והוא נמצא. הערכים שעליהם עוברים יודפסו עם פסיקים ביניהם. אם לא עוברים על אף מספר, מדפיסים שורה ריקה.
    - e. במידה ונקלט 4: הדפס את טבלת ה-hash לפי הפורמט המתאים.
- ניתן להניח שמספר האיברים בטבלת ה-hash במקרים של Hash סגור, לא יעלה על גודל הטבלה בשום שלב של ריצת התוכנית.
- ניתן להניח שלא יוכנס ערך שכבר קיים בטבלת ה-hash. מספר שלם בתרגיל זה הוא int.

## מימוש טבלאות ה-hash השונות.

### Chaining

- א. בהינתן גודל הטבלה  $L$ , פונקציית ה-hash היא  $h(x) = x \bmod L$ .
- ב. עליכם להוסיף איברים לתחילת הרשימה המקושרת.
- ג. פורמט ההדפסה של טבלת ה-hash הוא על פי הדפסה סידרתית של התאים כאשר פסיקים מפרידים בין כל שני תאים סמוכים. כל תא מודפס כך שהאינדקס מופיע ומיד אחריו מופיעים סוגריים שבתוכם מופיעים איברי הרשימה המקושרת מהתחלה לסוף. חובה לממש chaining על ידי מערך של רשימות מקושרות. מי שמימש בתרגיל הקודם גרפים לפי רשימת סמיכויות הרוויה.
- מי שלא יעשה כך – יקבל אפס בתרגיל כולו בלי זכות לערעור. ראו הוזהרתם!**

### Linear Probing עם הסרה לפי שיטת "צמצום רווחים"

- א. בהינתן גודל הטבלה  $L$ , פונקציית ה-hash היא  $h(x) = x \bmod L$ .
- ב. פורמט ההדפסה של טבלת ה-hash הוא פשוט הדפסת כל האיברים במערך כשפסיקים מפרידים ביניהם. בתאים ריקים לא ייכתב ערך.

### Double Hashing עם הסרה לפי שיטת המצבה

- א. בהינתן גודל הטבלה  $L$ , אנחנו מסמנים שני מספרים ראשוניים  $p_1$  ו- $p_2$  כך ש- $p_1$  הוא המספר הראשוני הקטן ביותר הגדול (ממש) מ- $L$ , ו- $p_2$  הוא המספר הראשוני הגדול ביותר שקטן (ממש) מ- $L$  (קיימים כאלה לכל  $L > 2$ ). למשל עבור  $L=8$  ייבחרו  $p_1=11$  ו- $p_2=7$ . עבור  $L=5$  ייבחרו  $p_1=7$  ו- $p_2=3$ .
- פונקציית ה-hash היא  $h(x, i) = (h_1(x) + i * h_2(x)) \bmod L$ , כאשר  $h_1(x) = x \bmod p_1$  ו- $h_2(x) = p_2 - (x \bmod p_2)$ .
- ב. פורמט ההדפסה של טבלת ה-hash הוא פשוט הדפסת כל האיברים במערך כשפסיקים מפרידים ביניהם. בתאים ריקים לא ייכתב ערך. מצבה תסומן על ידי  $X$ .
- ג. על מנת להקל עליכם, במקום כל מחיקה שימו מצבה.
- ד. במקרה בו לא ניתן למצוא תא ריק להכנסת איבר על ידי סדרת פונקציות ה-hash (עבור כל ערכי ה- $i$  השונים. שימו לב שיכול להיות מקרה כזה), אין להכניס את הערך כלל.

ראו דוגמאות בקובץ הפידבק.

### תזכורת נוספת:

**עליכם לבדוק היטב את התרגיל בסביבת ה-u2.**  
**בהוראת מרצה הקורס תהיה רק בדיקה אחת ולא יינתנו תיקונים, ולא משנה האם התרגיל עבד לכם בסביבות אחרות או שקיבלתם אפס על תקלה קטנה.**

בהצלחה!



זכרו! הכתובת לכל שאלה הנוגעת לתרגיל היא הפורום! בבקשה עבדו על פי הנהלים!