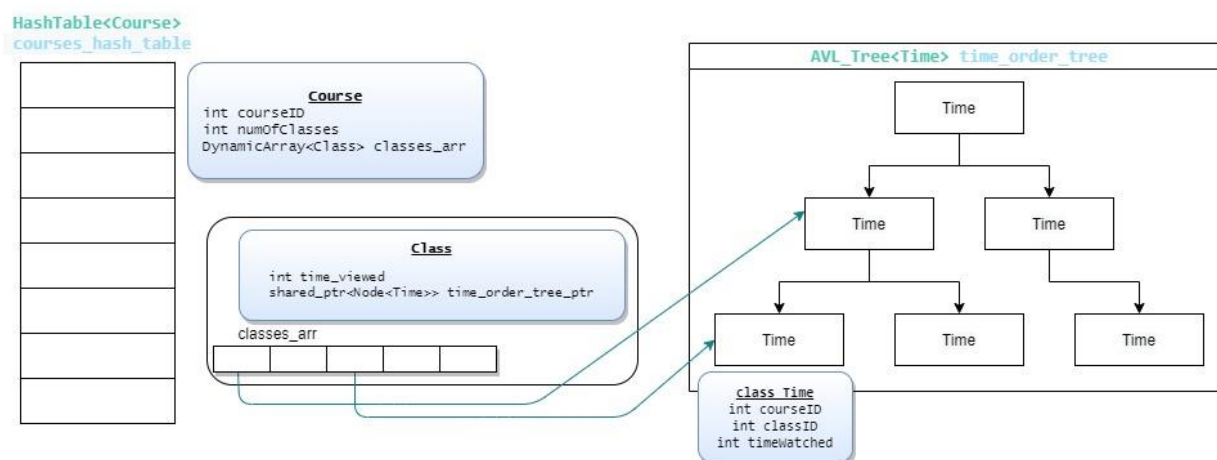


תרגיל רטוב 2 – חלק יבש

מגישות: דורין שטיימן ושני גורן



שרטוט סכמטי של מבנה הנתונים:

נבהיר כי החצים הכחולים מייצגים מצביעים, וכי המלבנים בצבע תכלת מייצגים מחלקות במבנה הנתונים.

המחלקות במבנה הנתונים:

Course

המחלקה שמייצגת קורס כפי שהוא נשמר בטבלת הערבול. מכילה את השדות הפנימיים:

- `courseID` – המספר המזהה של הקורס.
- `numOfClasses` – מספר השיעורים בקורס.
- `classes_arr` – מערך דינמי שכל תא בו מכיל `Class` – המייצג כל שיעור של הקורס הנ"ל השמור במערכת.

Class

המחלקה שמייצגת את השיעורים השמורים במערכת. מכילה את השדות הפנימיים:

- `time_viewed` – שומר את זמן הצפייה של כל שיעור. כל עוד לא הוספו דקות צפייה לשיעור, הערך מאותחל ל-0.
- `time_order_tree_ptr` – מצביע לצומת בעץ הצפיות בקורסים, המייצג את זמן הצפייה של השיעור. באיתחול, המצביע יאותחל ל-NULL, כי טרם הוספו צפיות לשיעור זה. בכל הוספה של זמן צפיה לקורס, עץ הצפיות יעודכן עם זמן הצפיה החדש, ולאחר מכן גם המצביע יעודכן להצביע על הצומת המתאים.

Time

המחלקה שמייצגת את הצפיות בשיעורים בעץ הצפיות. מכילה את השדות הפנימיים:

- `courseID` – המספר המזהה של הקורס.

- classID – המספר המזהה של השיעור.
- time_viewed – זמן הצפייה הכולל של השיעור.

הפעולות על מבנה הנתונים:

הערות:

1. נפרט את עיקרי האלגוריתם של כל פונקציה, תוך השמטה של בדיקות טריוויאליות של תקינות הקלט.
2. לשם מימוש מבנה הנתונים השתמשנו בטבלת ערבול דינאמית בשיטת chain hashing. פונקציית הערבול בה השתמשנו היא $f(k) = k \bmod(\text{size})$, כאשר size הוא הגודל הנוכחי של טבלת הערבול. ראינו בהרצאה כי פונקציה זו מקיימת את הנחת הפיזור האחד והפשוט, ולכן, אורך כל שרשרת יהיה $O(1)$ בממוצע על הקלט. מכאן נובע כי פעולת החיפוש של איברים בטבלת הערבול יהיו בסיבוכיות זמן $O(1)$ בממוצע על הקלט.
3. טבלת הערבול של הקורסים וכן מערך ההרצאות הנשמר עבור כל קורס מורכבים ממערכים דינאמיים. מערכים אלה מאוחלים בגודל 2, והם מוגדלים פי 2 כאשר הם מתמלאים, ומוקטנים פי 2 כאשר הם מגיעים לקיבולת של 25%. ראינו בתרגול כי סיבוכיות הזמן של פעולות הוספה ומחיקה של איבר ממערך דינמי כזה היא $O(1)$ משוערך. לכן בסך הכל, סיבוכיות הזמן של פעולות הוספה ומחיקה ממערך דינאמי היא $O(1)$ משוערך, וסיבוכיות הזמן של פעולות אלה עבור טבלת הערבול הדינאמית הן $O(1)$ משוערך בממוצע על הקלט.

Init()

אתחל את מבנה הנתונים: אתחל עץ AVL ריק, וטבלת ערבול ריקה בגודל 2.

סיבוכיות זמן: $O(1)$ – איתחול של עץ ריק דורש הקצאת כתובת לעץ והצבעה של השורש על NULL, ופעולות אלו מתבצעות ב- $O(1)$. עבור טבלת הערבול, נקצה שני תאים של קורס דיפולטיבי (מספר מזהה 1- וללא הרצאות) שכל אחד מהם מקצה מערך דינמי בגודל 2 של הרצאה דיפולטיבית (ללא זמני צפייה ופוינטר null לעץ הזמנים) ומכאן שסה"כ קורה בזמן פעולות לינארי $O(1)$.

סיבוכיות מקום: אתחול מבנה נתונים ריק כמתואר לעיל דרש מספר קבוע של הקצאות זכרון ב- $O(1)$.

AddCourse(void* DS, int courseID)

קרא לבנאי של Course עם ה-courseID.

חפש את הקורס המיועד להוספה בטבלת הערבול של הקורסים. במידה והקורס נמצא, זרוק כישלון וסיים. במידה והקורס לא נמצא, הוסף את הקורס לטבלת הערבול הדינאמית של כל הקורסים במערכת.

סיבוכיות זמן:

בניית איבר הקורס החדש: אתחול מערך דינמי בגודל 2 עבור ההרצאות העתידיות של הקורס (המערך יגדל בהתאם להכנסת הרצאות למערכת), אתחול numOfClasses ל-0 והשדה courseID עם המזהה שהוכנס למערכת עבור הקורס הנוכחי. כלומר, הבנאי יוצר עצם חדש מסוג קורס בזמן קבוע ב- $O(1)$.

חיפוש איבר בטבלת ערבול – לפי הערה 2, סיבוכיות הזמן של פעולה זו היא $O(1)$ בממוצע על הקלט.

הוספת הקורס לטבלת הערבול: לפי הערה 3, סיבוכיות הזמן של הוספת איבר לטבלת ערבול דינאמית הוא $O(1)$ משוערך בממוצע על הקלט.

שאר פעולות החישוב וההשוואה מבוצעות בזמן $O(1)$.

מכאן שסיבוכיות הזמן סה"כ: $O(1)$ משוערך בממוצע על הקלט.

סיבוכיות מקום:

הקצאת איבר מסוג Course חדש דורשת הקצאות בגודל קבוע ב- $O(1)$ מקום.

לשם מימוש הפונקציה נעזרנו בפעולת הוספת הקורס לטבלת ערבול, פעולה זו מקצה אך ורק משתנים מקומיים מטיפוסים פשוטים ב- $O(1)$ מקום.

כפי שראינו בתרגול, אילו הקורס שהוספנו גורר צורך בהגדלת המערך (נוסף בתא הריק האחרון), יש להקצות מערך גדול פי 2 וללערבל מחדש את איברי המערך הנוכחי לתוכו. הקצאת המערך החדש עומד בסיבוכיות מקום $O(2n) = O(n)$ ומעבר לכך אין צורך בהקצאות נוספות לשם ההגדלה.

סה"כ סיבוכיות המקום: $O(1) + O(1) + O(n) = O(n)$.

RemoveCourse(void* DS, int courseId)

מצא את הקורס המיועד למחיקה בטבלת הערבול של הקורסים. במידה והקורס לא נמצא, זרוק כישלון וסיים.

עבור על מערך ההרצאות הדינמי של הקורס הנוכחי. גש באמצעות הפוינטר של ההרצאה לצומת המתאים בעץ הצפיות. במידה וקיים, מחק את הצומת הנ"ל.

לבסוף, מחק את הקורס מטבלת הערבול הדינמית של הקורסים השמורים במערכת.

סיבוכיות זמן:

חיפוש הקורס בטבלת הערבול: לפי הערה 2, סיבוכיות הזמן של חיפוש בטבלת ערבול הוא $O(1)$ בממוצע על הקלט.

מחיקת הצפיות מעץ הצפיות – בהינתן שמערך ההרצאות של הקורס הוא בגודל m . בכל איטרציה במעבר על המערך, מבצעים מחיקה מעץ הצפיות, שבו יש M צמתים לכל היותר (אילו צפו בכל הרצאה בכל קורס). כפי שראינו בהרצאה, סיבוכיות הזמן של מחיקת צומת מעץ AVL היא $O(\log(M))$. לכן הסיבוכיות תהיה $O(m \log(M))$.

מחיקת הקורס מטבלת הערבול הדינמית: לפי הערה 3, סיבוכיות הזמן של מחיקת איבר מטבלת ערבול דינמית היא $O(1)$ בממוצע על הקלט משוערך.

שאר פעולות החישוב וההשוואה מבוצעות בזמן $O(1)$.

סיבוכיות הזמן הכוללת של הפונקציה:

$$O(m \log(M)) = O(m \log(M)) + O(1) + O(1) \text{ בממוצע על הקלט, משוערך.}$$

סיבוכיות מקום:

לשם מימוש הפונקציה נעזרנו בפעולת מחיקת צומת מעץ AVL (עץ הצפיות) ובה מתבצעות שתי הקצאות:

מערך לשמירת המסלול אל הצומת שהוסר בעץ הצפיות, באורך גובה השורש. מההרצאה, ידוע כי בעץ AVL $|h_{root}| = O(\log(M))$, כאשר m הוא מספר הצמתים בעץ.

מחסנית הקריאות של פונקציית החיפוש בעץ הצפיות. סיבוכיות הזמן של החיפוש היא, כפי שהראנו, $O(\log(M))$. בכל קריאה רקורסיבית נבצע פעולות השוואה ב- $O(1)$ ולכן סיבוכיות המקום של פונקציית החיפוש היא: $O(\log(M))$.

כפי שראינו בתרגול, אילו הקורס שהוספנו גורר צורך בהקטנת המערך (אחרי מחיקתו קיבולת המערך הדינמי הינה 25%), יש להקצות מערך קטן פי 2 ולערבל מחדש את איברי המערך הנוכחי לתוכו. הקצאת המערך החדש עומד בסיבוכיות מקום $O\left(\frac{n}{2}\right) = O(n)$ ומעבר לכך אין צורך בהקצאות נוספות לשם ההקטנה.

סה"כ סיבוכיות מקום: $O(\log(M)) + O(\log(M)) + O(n) = O(\log(M) + n)$.

AddClass

גש לקורס courseID בטבלת הערובול של הקורסים. במידה והקורס לא נמצא זרוק כשלון וסיים. אתחל איבר מסוג Class, והוסף אותו למערך הדינמי של השיעורים בקורס זה.

ההוספה למערך הדינמי מתבצעת כך: נשמר מונה של מספר האיברים במערך הדינמי, והאיבר האחרון שנוסף למערך מוכנס לאינדקס של מספר המונה. לפני ההכנסה מתבצעת בדיקה האם המערך מלא. במידה וכן, מתבצעת הגדלה של המערך פי 2.

לאחר ההכנסה למערך, הגדל את מספר השיעורים בקורס ב-1, והחזר את המספר המזהה של השיעור שנוסף.

סיבוכיות זמן: גישה לאיבר בטבלת ערובול – לפי הערה 2, סיבוכיות הזמן של פעולה זו היא $O(1)$ בממוצע על הקלט.

איתחול איבר מסוג Class - $O(1)$.

הוספת איבר למערך דינאמי – לפי הערה 3, סיבוכיות הזמן המשוערכת של הוספת איבר למערך דינאמי היא $O(1)$. סיבוכיות משוערכת זו כוללת בתוכה גם את פעולת הגדלת המערך המתבצעת במידת הצורך.

סיבוכיות מקום: הקצאת איבר מסוג Class - $O(1)$.

במקרה הגרוע, מתבצעת הקצאת זיכרון עבור מערך דינאמי חדש של השיעורים בקורס. כפי שראינו בתרגול, הזיכרון שמוקצה במקרה זה הוא לכל היותר פי 2 ממספר השיעורים השמורים עבור הקורס אליו התווסף השיעור. מספר השיעורים בקורס מסוים חסום מלמעלה על ידי m – מספר השיעורים הכולל בכל מבנה הנתונים. לכן המקום המוקצה הוא לכל היותר $O(m) = 2m$, ומכאן שבמקרה הגרוע סיבוכיות המקום של הפונקציה תהיה $O(m)$.

WatchClass

גש לקורס courseID בטבלת הערובול של הקורסים. במידה והקורס לא נמצא זרוק כשלון וסיים.

במידה והקורס נמצא, בדוק האם המספר המזהה של ההרצאה לה אנו מעוניינים להוסיף זמן צפייה נמצא בטווח מספר ההרצאות של הקורס courseID. במידה ולא, זרוק שגיאת קלט לא חוקי וסיים.

במידה והמספר המזהה של ההרצאה חוקי, עדכן את זמן הצפייה של ההרצאה במערך ההרצאות של הקורס הנוכחי ע"י הוספת time לזמן הצפייה של ההרצאה.

גש לתא המייצג את ההרצאה המבוקשת במערך ההרצאות של הקורס הנוכחי, וגש למצביע לצומת בעץ הצפיות השמור בתא זה. אם הוא מצביע על צומת בעץ (ולא על NULL), מחק מעץ הצפיות את האיבר של זמן הצפייה של ההרצאה classID של קורס courseID.

צור איבר צפיות חדש מסוג Time לעץ הצפיות עבור קורס courseID והרצאה classID עם זמן הצפייה המעודכן, והוסף איבר זה לעץ הצפיות.

עדכן את המצביע לאיבר החדש בעץ הצפיות במערך ההרצאות של הקורס הנוכחי באינדקס ה-classID.

סיבוכיות זמן:

גישה לאיבר בטבלת ערבול – לפי הערה 2, סיבוכיות הזמן של פעולה זו היא $O(1)$ בממוצע על הקלט. איתחול Class - $O(1)$.

גישה לאיבר במערך דינאמי - $O(1)$.

מחיקת איבר הצפיה הישן והוספת האיבר המעודכן לעץ הצפיות בקורסים: בעץ הצפיות יש M צמתים (במקרה הגרוע ביותר, לכל הרצאה בכל קורס יש צפייה - כלומר, מהווה איבר בעץ הצפיות. סך כל ההרצאות במבנה מיוצג על ידי M). סיבוכיות הזמן של המחיקה וההוספה היא $O(\log(M)) = 2 \log(M)$.

לכן בסך הכל סיבוכיות הזמן של הפונקציה היא $O(\log(M))$ בממוצע על הקלט.

סיבוכיות מקום:

הקצאת איבר מסוג Time - הקצאת הזכרון ל-int time, int courseID, int classID היא $O(1)$.

לשם מימוש הפונקציה נעזרנו בפונקציות החיפוש והמחיקה של עץ AVL. כפי שהוסבר קודם, סיבוכיות המקום של כל אחת מהן היא $O(\log(M))$.

סה"כ סיבוכיות מקום: $O(\log(M))$.

TimeViewed

גש לקורס courseID בטבלת הערבול של הקורסים. במידה והקורס לא נמצא זרוק כשולן וסיים.

במידה והקורס נמצא, בדוק האם ההרצאה classID בטווח מספר ההרצאות של הקורס courseID, במידה ולא, זרוק שגיאת קלט לא חוקי וסיים.

במידה וההרצאה classID חוקית, גש לתא במערך ההרצאות המייצג את ההרצאה המבוקשת, והחזר את הערך ששמור בשדה time_viewed של הרצאה זו.

סיבוכיות זמן:

גישה לאיבר בטבלת ערבול – לפי הערה 2, סיבוכיות הזמן של פעולה זו היא $O(1)$ בממוצע על הקלט. גישה לאיבר במערך דינאמי - $O(1)$.

מכאן שסיבוכיות הזמן הכוללת של הפונקציה היא $O(1)$ בממוצע על הקלט.

סיבוכיות מקום:

בפונקציה זו לא מתבצעות הקצאות מקום או קריאות רקורסיביות, ולכן סיבוכיות המקום שלה היא $O(1)$.

GetIthWatchedClass

גש לעץ הצפיות, ושמור את המצביע לצומת המתקבל מפעולת select(i) על עץ הדרגות. פעולה זו מומשה על ידי האלגוריתם שנלמד בתרגול. אם מפעולת ה-select לא הוחזר מצביע, החזר כישולן וסיים. (פעולת ה-select תחזיר NULL אם לא קיימים מספיק צמתים בעץ הדרגות).

אחרת, גש לשדות הפנימיים של הצומת שהוחזר, השומרים את המספרים המזהים של הקורס וההרצאה, ושמור את ערכיהם במצביעים courseID, classID שנשלחו לפונקציה כפרמטרים.

סיבוכיות זמן:

Select בעץ דרגות – כפי שלמדנו בתרגול, סיבוכיות הזמן של פעולה זו היא $O(\log(M))$, כאשר M הוא מספר הצמתים בעץ.

גישה ועדכון של מצביעים ושדות פנימיים - $O(1)$.

לכן, סיבוכיות הזמן הכוללת של הפונקציה היא $O(\log(M))$ במקרה הגרוע.

סיבוכיות מקום:

המימוש של הפונקציה select על עץ דרגות הוא רקורסיבי, וכפי שלמדנו בתרגול, עומק הרקורסיה של פעולה זו היא $O(\log(M))$. בכל קריאה רקורסיבית נבצע פעולות השוואה ב- $O(1)$ ולכן סיבוכיות המקום של הפונקציה היא: $O(\log(M))$.

פרט לכך לא מתבצעות הקצאות נוספות, ולכן סיבוכיות המקום הכוללת היא $O(\log(M))$.

Quit()

הפונקציה תמחק את העץ הצפיית על ידי כך שתעבור על הצמתים בעץ זה ותקרא להורס המתאים של הטיפוס Time השמור בכל צומת בעץ.

כמו כן, תבצע מחיקה של טבלת הערבול של הקורסים. בעת המחיקה יקרא ההורס של הרשימה המקושרת השמורה בכל תא במערך (טבלת הערבול מומשה בעזרת chain hashing) ובה שמורים איברים מסוג Course. הורס הרשימה המקושרת יקרא להורס של הטיפוס Course לשם מחיקת האיברים. נזכור כי בקורס שמור מערך דינמי של ההרצאות ולכן יקרא גם הורסו שבעת מחיקת כל איבר יקרא להורס של הטיפוס Class.

סיבוכיות זמן:

מחיקת עץ הצפיית: באיבר מסוג Time שלושה משתנים בודדים שנשחררם ב- $O(1)$. בעץ הנ"ל לכל היותר m צמתים כמספר כל ההרצאות במבנה (אילו לכל הרצאה נוספה לפחות צפייה אחת). מכאן ששחרור כל כתובות הזכרון לשם מחיקת העץ יקח $O(m)$ בסיוור inorder.

מחיקת טבלת הערבול הדינמית של הקורסים: כפי שתואר, קריאות ההורס העוקבות לשם מחיקת כתובת זכרון במערך טבלת הערבול הן: קריאה להורס של הרשימה המקושרת מטיפוס $course \leftarrow$ קריאה להורס של הטיפוס Course \leftarrow קריאה להורס של מערך דינמי מטיפוס Class \leftarrow קריאה להורס של הטיפוס Class.

*נדגיש כי שאר המחיקות המתבצעות הן של משתנים בודדים של הטיפוסים השונים ומתבצעות ב- $O(1)$.

- נשים לב כי מחיקת טיפוס מסוג Class היא $O(1)$ כי הוא מחזיק רק שני משתנים פשוטים. מחיקת המערך דינמי תקח $O(num_of_classes) = O(size)$ כאשר $size$ גודל מערך ההרצאות. זאת משום שהמערך גדל וקטן בפרופורציה למספר התאים הלא ריקים (לא יחזיק פחות איברים מרבע מהגודל שלו). מכאן שמחיקת כל איברי הטיפוס מסוג Class במערכים הדינאמיים של הקורסים השונים יקח זמן פרופורציונלי לכמותם במערכת (מוסמן ב- m) כלומר $O(m)$.

- טבלת הערבול היא דינמית גם כן ולכן קטנה וגדלה בפרופורציה למספר הרשימות המקושרות הלא ריקות בטבלה (כאמור, לא יחזיק בשום מצב נתון פחות מ- $size_of_hash_table/4$ רשימות לא ריקות מהגודל שלו) ומכאן שבהכרח קיים $c > 0 \exists N$ כך ש: $size_of_hash_table < c \cdot n$.

עבור כל הקורסים יחדיו ימחק מידע כולל נוסף מסדר גודל של $O(m)$ – מספר ההרצאות במערכת.

לכן סה"כ סיבוכיות זמן למחיקת טבלת הערבול הדינמית של הקורסים: $O(n + m)$.

וסיבוכיות הזמן הכוללת: $O(m + n) + O(m) = O(m + n)$.

סיבוכיות מקום:

הפונקציה לא משתמשת במקום נוסף מלבד במחסנית הקריאות עבור מעבר על צמתי העץ ב-Inorder. כפי שראינו בהצגה, עומק עץ הצפיות הוא לכל היותר $\log(m)$ כאשר m מספר כל ההרצאות במערכת וכך גם עומק הרקורסיה (לא נרד יותר מעבר לעלה הנמוך ביותר). סיבוכיות המקום לפעולה אם כן היא: $O(\log(m))$.

סיבוכיות המקום עבור כל מבנה הנתונים:

שמירת עץ הצפיות (עץ AVL) דורשת לכל היותר $O(m)$ מקום כמספר ההרצאות הקיימות במערכת (בהנחה ולכל הרצאה זמן צפייה גדול מ-0). טבלת הערבול הדינמית עבור כלל הקורסים במערכת הינה מסדר גודל של n וכלל הקורסים מחזיקים מערכים דינמיים בסדר גודל של `num_of_classes` (פירוט מורחב בהסבר הפעולה Quit) של הקורס הנוכחי ואורכם הכולל מסתכם אם כך לסדר גודל של m , כמספר כלל ההרצאות במערכת. סה"כ, החזקת המערך הדינמי דורשת סיבוכיות מקום בסדר גודל של $O(m + n)$.

כמו כן, סיבוכיות המקום הדרושה לביצוע של כל הפונקציות של מבנה הנתונים חסומות מלעיל על ידי $O(m + n)$. לכן בסך הכל, סיבוכיות המקום של המבנה, שהיא סך כל הזיכרון המוקצה וסיבוכיות הזמן של הפעולות עליו, היא $O(m + n) + O(m) = O(m + n)$.