

Insert

v - צומת חדש
v - צריך להיות אבא של v
g אבא של p (סבא של v)
הכנס v כבן נוסף של p
- אם ל p יש שלושה ילדים - גמרת.
- אם ל p יש ארבעה ילדים:
פצל ל - p שמאל ו- p' מימין ושתול
את p' כבן של g

Delete

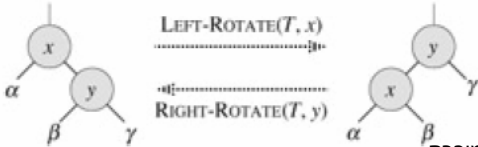
v - צומת, p - אב, u - דוד.
- השמט v מ p
(1) אם ל v שני אחים: השמט וסיים.
(2) אם ל v א בודד:
א) אם ל u שלושה בנים:
p יאמץ אחד מבני הדוד
ב) אם ל u שני בנים:
u יאמץ את האח של v ועכשיו p
ערירי ויש להשמיטו רקורסיבית

ייצוג קבוצות יעיל התומך בפעולות איחוד (זר) ו-Find ו-Insert.

האיברים ממוספרים 1, 2, ... והקבוצות A, B.
1. נשמור מערך ובו בתא i נשמור את הקבוצה I משתייך.
Find, Insert: O(1) Merge: O(n)
2. נשמור לכל קבוצה את גודלה ואת האיבר הראשון בה ולכל איבר נשמור את הקבוצה אליה הוא משתייך ואת האיבר הבא אחריו.
Merge: עוברים על איברי הקבוצה הקטנה ומשנים את הקבוצה שלהם לגדולה, את האיבר האחרון משרשרים לאיבר הראשון בקבוצה הגדולה ומשנים את מאפייני הקבוצה הגדולה בהתאם. מאחר וכל איבר שעובר לבעלים חדשים מגדיל לפחות פי 2 את הבעלים, כל איבר עובר לכל היותר lg(n) פעמים במהלך n פעולות Merge וכן הסיבוכיות לח פעולות כאלו היא O(n * lg(n)).
3. נשמור על עבר כל קבוצה בו כל צומת מצביע לאביו (שהוא האיבר הבא אחריו) ובשורש העץ נמצא שם הקבוצה.
Merge: תולים שורש של אחת הקבוצות על השנייה. O(1).
Find: רצים כלפי מעלה. O(n).
אך אם נקפיד לתלות עץ קטן על עץ גדול נגדיל את עומק העץ ב-1 בכל תלייה ובכל תלייה מספר האיברים בקבוצה לפחות מוכפל. הסיבוכיות לח פעולות Find היא O(n * lg(n)).

עצים אדומים שחורים

עץ חיפוש בינארי "כמעט מאוזן" בו לכל צומת יש צבע שחור או אדום. ערכו של כל עלה הוא Null וצבעו שחור. לכל צומת אדום שני ילדים שחורים. כל מסלול מצומת מסוים לכל עלה הוא בעל אותו מספר של צמתים שחורים. bh(T) הוא הגובה השחור של עץ- מספר הצמתים השחורים מהשורש לעלה (לא כולל השורש).
משפט: גובהו של עץ אדום-שחור בעל n צמתים פנימיים הוא לכל היותר 2lg(n+1).
רוטציה: פעולת סיבוב השומרת על הסדר הנתוני (Inorder) של העץ. כאשר מפעילים רוטציה שמאלית של צומת x, אנו מניחים שהבן הימני שלו y הפכת את y לשורש החדש של תת-העץ, את x לבן השמאלי של y ואת הבן השמאלי של y לבן הימני של x.
זמן ריצה: O(1).



RB-INSERT(T, z)
Tree-Insert(T, z)
Color[z] ← Red
1 while z ≠ root and color[p[z]] = RED
2 do if p[z] = left[p[p[z]]]
3 then y ← right[p[p[z]]]
4 if color[y] = RED
5 then color[p[z]] ← BLACK // Case 1
6 color[y] ← BLACK // Case 1
7 color[p[p[z]]] ← RED // Case 1
8 z ← p[p[z]] // Case 1
9 else if z = right[p[p[z]]]
10 then z ← p[z] // Case 2
11 LEFT-ROTATE(T, z) // C.2
12 color[p[z]] ← BLACK // Case 3
13 color[p[p[z]]] ← RED // Case 3
14 RIGHT-ROTATE(T, p[p[z]]) // C.3
15 else (same as then clause with "right" and "left" exchanged)
16 color[root[T]] ← BLACK

Maintaining Dynamic Tree with Min Subject to Constraints

כאשר רוצים לבצע שאילתות לפי שני מפתחות (למשל ציון וגובה) נשמור שני עצי-3, בכל אחד מהם נשמור מפתחות שונים (באחד גבהים ובשני ציונים) ונשמור מצביעים בין הרשומות (בין הציון של תלמיד לגובהו שלו למשל, נניח כי הציונים ייחודיים). כשנרצה לחפש למשל את הציון המינימלי של התלמידים החל מגובה מסוים נסרוק את תת-העץ של התלמידים מהגובה הרצוי (קל למצוא) ומציץ כלפי מעלה את הציון המינימלי שלהם. כל פעולה לוקחת O(lg(n)).

Splay Trees
עץ חיפוש בינארי בו פעולה בודדת לוקחת O(n) אבל m פעולות לוקחות O(m lg(n)). מבנה זה יעיל לשימוש כאשר הסיכויים לגישה חוזרת גבוהים.

גידול של פונקציות

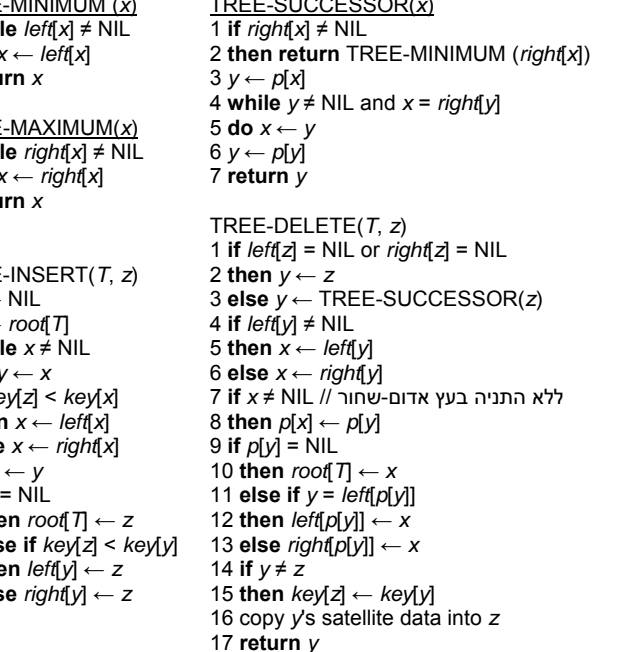
We say that f(n) has order of growth O(g(n)) if there are constants c1 ≠ 0 and c2 ≠ 0 such that for all n > 0 c1 * g(n) ≤ f(n) ≤ c2 * g(n)
במקרה זה נאמר כי g(n) הוא חסם הדוק אסימפטוטי עבור f(n).
We say that T(n) ∈ O(f(n)) if there is a constant c ≠ 0 such that for all n > 0 T(n) ≤ c * f(n)
במקרה זה נאמר כי g(n) הוא חסם אסימפטוטי עליון עבור f(n).
We say that T(n) ∈ Ω(f(n)) if there is a constant c ≠ 0 such that for all n > 0 c * f(n) ≤ T(n)
במקרה זה נאמר כי g(n) הוא חסם אסימפטוטי תחתון עבור f(n).
הסימון o(f(n)) משמש להגדרת חסם עליון שאינו הדוק אסימפטוטי.
הסימון ω(f(n)) משמש להגדרת חסם תחתון שאינו הדוק אסימפטוטי.
שיטות לפתרון נוסחאות נסיגה: הצבה - חישוב ובדיקה, ניתן גם עם קבוע, למצוא הדוק ביותר, איטרציה, עץ רקורסיבי, שיטת אב.

מבני נתונים

טבלאות דינאמיות
מערך עם פעולות הכנסה ומחיקה. כאשר התמלא השטח המוקצה, מוקצה שטח חדש (בד"כ כפול בגודלו), מועתקים אליו הערכים הישנים ושאר התאים מאותחלים. עבור n פעולות הכנסה, סדר הגודל הוא 3n ולכן עבור פעולה יחידה סדר הגודל הוא 3 (Amortized).
עצים - כלי
אורך מסלול - מספר הקשתות בין זוג צמתים.
גובה העץ - אורך המסלול הגדול ביותר בין שורש העץ לאחד העלים.
עומק של צומת - אורך המסלול ממנה לשורש העץ.
אם לא הוגדר אחרת, לכל צומת בעץ יש מצביע להורה שלו (Null בשורש).
סריקות עצים
DLR - Preorder: בקר בשורש, סרוק את תת-עץ שמאלי, סרוק תת-עץ ימני.
LDR - Inorder: סרוק תת-עץ שמאלי, בקר בשורש, סרוק תת-עץ הימני.
LRD - Postorder: סרוק תת-עץ שמאלי, סרוק תת-עץ ימני, בקר בשורש.
יימוש עצים
באמצעות מערך הורים - מעניקים לכל צומת מספר ומחזיקים שני מערכים. במערך הראשון ב A[i] נמצא המספר של ההורה של i ובמערך השני ב B[i] נמצא הערך של i (הנתונים). בעייתיות: לא ניתן לדעת את סדר הבנים (משמאל לימין או מימין לשמאל) ולכן יש להגדיר זאת מראש.
ע"י רשימות בנים - מחזיקים מערך עבור כל הצמתים בעץ כך שכל תא במערך מצביע לרשימה (רצוי בהקצאה דינאמית) של הבנים שלו.
ע"י מערך סמנים - לכל תא במערך יש ערך וכתובות של הבן הימני והשמאלי במערך (Null אם אין כאלו).
בהקצאה דינאמית - ידוע.
עץ חיפוש בינארי
לכל תת-עץ, כל הערכים בתת-העץ השמאלי קטנים מאלו של השורש וכל הערכים בתת העץ הימני גדולים ממנו.
מבנה זה רלוונטי רק כאשר יש יחס סדר בין האיברים.
סריקת Inorder של עץ כזה תציג רשימה ממוינת של האיברים בו.

TREE-MINIMUM(x)
1 while left[x] ≠ NIL
2 do x ← left[x]
3 return x
TREE-SUCCESSOR(x)
1 if right[x] ≠ NIL
2 then return TREE-MINIMUM(right[x])
3 y ← p[x]
4 while y ≠ NIL and x = right[y]
5 do x ← y
6 y ← p[y]
7 return y
TREE-MAXIMUM(x)
1 while right[x] ≠ NIL
2 do x ← right[x]
3 return x
TREE-DELETE(T, z)
1 if left[z] = NIL or right[z] = NIL
2 then y ← z
3 else y ← TREE-SUCCESSOR(z)
4 if left[y] ≠ NIL
5 then x ← left[y]
6 else x ← right[y]
ללא התניה בעץ אדום-שחור // if x ≠ NIL
8 then p[x] ← p[y]
9 if p[y] = NIL
10 then root[T] ← x
11 else if y = left[p[y]]
12 then left[p[y]] ← x
13 else right[p[y]] ← x
14 if y ≠ z
15 then key[z] ← key[y]
16 copy y's satellite data into z
17 return y

עצי 2-3
לכל צומת פנימי 2-3 ילדים, לכל העלים אותו עומק. האלמנטים הנשמרים בעלים מסודרים בסדר עולה משמאל לימין. בכל צומת פנימי שומרים את ערך המפתח הנמוך ביותר בתת-העץ האמצעי ואת הנמוך ביותר בתת-העץ הימני (אם קיים).



עצי 2-3
לכל צומת פנימי 2-3 ילדים, לכל העלים אותו עומק. האלמנטים הנשמרים בעלים מסודרים בסדר עולה משמאל לימין. בכל צומת פנימי שומרים את ערך המפתח הנמוך ביותר בתת-העץ האמצעי ואת הנמוך ביותר בתת-העץ הימני (אם קיים).

ערימה מערך המייצג עץ בינארי "כמעט" שלם: העץ מלא לחלוטין פרט אולי לרמתו האחרונה, המלאה משמאל ועד לנקודה מסוימת. דיועים גודל המערך וגודל הערימה. הערימה: Parent(i)=i/2, Left(i)=2i, Right(i)=2i+1. הערך בכל צומת קטן או שווה מהערך של ההורה.

Heapify פעולה המקבלת מערך וצומת i כך שהבנים של i הם ערימות חוקיות ויוצרת ערימה חוקית בצורה רקורסיבית כלפי מטה O(lg(n)).
בניית ערימה - סורקים חצי המערך העליון ומבצעים Heapify. זמן הריצה הוא לכאורה O(n lg(n)) אך ניתן למצוא חסם הדוק יותר והוא O(n).
Extract-Max - מחזיקים השרש, במקומו שמים את הקטן ביותר מבצעים heapify.
שיטת מין - מודל ההשוואות במודל ההשוואות מתארים את המינין כעץ החלטה לפי יחס הסדר בין האיברים. העלים בעץ הם הסדרים האפשריים (עבור n איברים יהיו n! סדרים - עלים). **משפט**: גובהו של עץ החלטה המינין n איברים הוא O(n lg(n)), כלומר לא ניתן למיין בזמן ריצה טוב יותר מ-O(n lg(n)), אלא עד כדי קבוע. הנקחה: עץ בינארי מכליל לולאת היתור 2^n עלים, מכאן n! <= 2^n n! ומכאן lg(n!) >= n lg(n) ומאחר ו-O(n lg(n)) >= n lg(n/2) >= n lg(n) >= n lg(n/2) >= n lg(n).
 לפי מודל זה, החסם התחתון לעץ השוואות בעל n עלים הוא O(lg(n)).

מין הכנסה - O(n^2)
מין ערימה - O(n lg(n))
HEAPSORT(A) בכל איטרציה i האיברים i...1 ממויינים
 1 BUILD-MAX-HEAP(A) ע"י החלפות.
 2 **for** i ← length[A] **downto** 2 **do** exchange A[i] ↔ A[1] **מיון בחירה** - O(n^2)
 3 **do** exchange A[1] ↔ A[i] בכל איטרציה i שמים במקום ה-i את האיבר הקטן מבין האיברים ח...1.
 4 **heap-size**[A] ← **heap-size**[A]-1
 5 MAX-HEAPIFY(A, 1)

מיון מיוזג - O(n lg(n)) מחלקים את המערך לשני חלקים שווים, ממינים כל חלק רקורסיבית ומזגים.
מיון מהיר - O(n lg(n)) Avg. O(n^2) WC - בוחרים Pivot מחלקים את המערך לערכים גדולים ממנו וקטנים ממנו (Partition) וממשיכים רקורסיבית. Partition(i,j) לוקח O(j-i+1) מאחר ומבקרים בכל איבר פעם אחת. המקרה הגרוע הוא זה בו מחלקים את המערך לחלקים בגודל 1 ו-(n-1): O(n^2) ∈ T(n) = T(n-1) + O(n) המקרה האידיאלי הוא בו מחלקים את המערך תמיד לשני חלקים שווים: O(n lg(n)) ∈ T(n) = 2T(n/2) + O(n). כל Partition ביחס קבוע (למשל 1:9) ייתן זמן ריצה של O(n lg(n)), זהו המקרה הממוצע.

שיטת מיון נוספות
מיון מנייה (Count) - O(n+k)
 אלגוריתם למיין n מספרים בתחום 1..k. למיין מערך A בגודל k נסרוק אותו ונשמור במערך B בגודל k את מספר המופעים של כל איבר בתחום. נסרוק את C ונצבור בכל תא את מספר האיברים הקטנים ממנו. נסרוק שוב את A ועבר כל איבר נניס למערך B בגודל n את האיבר Am לפי המיקום עליו מורה C ונפחית 1 מ-C.
מיון בסיס (Radix) - O(d*n)
 ממיין n מספרים בני d (קבוע) ספרות כל אחד. בכל שלב ממוינים המספרים בשיטת מיון "ציבה" (למשל Count Sort) לפי ספרות החל LSD. דוגמה למיין n איברים בתחום 1-n^2: מספר הספרות של המספרים בתחום הוא לכל היותר lg(n^2) = 2 lg n, יניצג כל מספר כזה במערך בגודל 2 ונבצע Radix עם שימוש ב CountSort. סה"כ: O(n+n) = Bin Sort / Bucket Sort
מיון מנייה (Count) - O(n+k) * O(n+n) = Bin Sort / Bucket Sort

מחלקים את הקלט במערך בן n האיברים ל-n קבוצות. נעזרים במערך B של n רשימות מקושרות וסורקים את A כדי להכניס כל איבר בו לרשימה המתאימה (הקבוצה) ב-B. לבסוף ממינים את רשימות B בעזרת מיון הכנסה. טוב למחזורות באורך קבוע או תחומים שידוע עליהם אינפורמציה, למשל: 1..n.
מציאת statistic - O(n) worst case
 בחר אלמנט i מתוך n איברים: (select)
 (1) חלק האיברים ל-5 קבוצות של 5 איברים
 (2) מצא את החציון של כל מהקבוצות (מיון פשוט)
 (3) השתמש רקורסיבית ב select למציאת חציון החציונים
 (4) בצע partition מסביב לחציון החציונים
 יהי א מספר האיברים בחלק הנמוך
 יהי א-k מספר האיברים בחלק הגבוה
 (5) השתמש ב select רקורסיבית למציאת i בחלק הנמוך
 או א-i בחלק הגבוה

Order Statistics
 ככלל, נוכל למיין מערך ולמצוא את האיבר ה-i או ב-O(n lg(n)).
 שיטה נוספת היא לחלק את המערך באמצעות Partition אקראי ולחפש את האיבר בחלק המתאים. בתוחלת O(n) WC, O(n2).
טבלאות Hash
 מבנה נתונים טוב לייצוג קבוצות, בממוצע זמן קבוע, תחליף לוקטור ביטים כאשר התחום גדול.

הנותן Hash
 מחזיקים מערך בן B "סלים" (תאים) כך שכל תא מצביע לרשימה מקושרת ממויינת. הגישה לכל סל היא O(1) והחיפוש ברשימה הוא לינארי. **פונקציית Hash** h(x) היא פונקציה המתאימה לכל ערך x את הסל המתאים לו בין 0 לבין (B-1).
 עבור N איברים ו-B סלים יהיו בממוצע N/B (מקדם העומס-α) איברים בכל סל ועבור פונקציה שמפזרת בצורה אחידה (ההסתברות של מפתח אקראי ליפול בכל תא היא 1/B) חיפוש (כושל או מצליח) אורך בממוצע O(1+α) עבור חישוב הפונקציה α-1 עבור סריקת הרשימה). אם נבחר B ביחס טוב למספר האיברים: N = O(B) אזי α=1 ואז פעולת החיפוש לוקחת O(1). דוגמאות לפונקציות: עבור 0 <= k < 1: h(k) = [kB], עבור k שלם: h(k) = k mod B (רצוי B ראשוני).

מערך Hash
 מחזיקים מערך בן B "סלים" ללא רשימות מקושרות. אם הפונקציה גרמה להתנגשות מחשבים פונקציות נוספת h2(x), h(x). למשל: H(x) = (H(x)+i) mod B
 בחיפוש איבר נסרוק את התאים לפי h1(x), h2(x) ועד שנגיע למקום ריק או נדע שהאיבר לא נמצא. (יש לשמור ערך עבור "ריק" ו-"deleted").

Hash אוניברסלי
 אם רוצים למנוע ממשתמש "דדוני" לחבל ביעילות, נבחר בכל פעם מחזיקים מערך בן B "סלים" ללא רשימות מקושרות. אם הפונקציה גרמה להתנגשות מחשבים פונקציות נוספת h2(x), h(x). למשל: H(x) = (H(x)+i) mod B
 בחיפוש איבר נסרוק את התאים לפי h1(x), h2(x) ועד שנגיע למקום ריק או נדע שהאיבר לא נמצא. (יש לשמור ערך עבור "ריק" ו-"deleted").

Hash אוניברסלי
 אם רוצים למנוע ממשתמש "דדוני" לחבל ביעילות, נבחר בכל פעם מחזיקים מערך בן B "סלים" ללא רשימות מקושרות. אם הפונקציה גרמה להתנגשות מחשבים פונקציות נוספת h2(x), h(x). למשל: H(x) = (H(x)+i) mod B
 בחיפוש איבר נסרוק את התאים לפי h1(x), h2(x) ועד שנגיע למקום ריק או נדע שהאיבר לא נמצא. (יש לשמור ערך עבור "ריק" ו-"deleted").

ובאופן אקראי את הפונקציה שלנו מתוך קבוצה אוניברסלית H- קבוצה המבטיחה לנו: H = { f | f: U -> {0,..., B-1} }
 ולכל x,y ∈ U מתקיים |h(x)-h(y)| ≤ |H|/B
 כלומר אם פונקציה נבחרת באופן אקראי מH אוניברסלית אז הסיכוי להתנגשות של x ו-y הוא 1/B. **קבוצה 2-אוניברסלית**: לכל 2 מפתחות שונים x,y ופונקציה h הנבחרת נדומלית מתוך H אז הסדרה <h(x),h(y)> היא כל אחת מ־m הסדרות האפשריות בהסתברות שווה. כל קבוצה 2-אוניברסלית היא אוניברסלית. לא כל קבוצה אוניברסלית היא 2-אוניברסלית.

Expected Average Case	Expected Worst Case	Open Hash
O(1)	O(logn)	Search(Key)
O(1)	O(logn)	Insert(Key)
O(1)	O(logn)	Delete(Key)
O(n)	O(n)	Minimum()
O(n)	O(n)	Maximum()

משפט: אם h נבחרת מאופן אוניברסלי של פונקציות hash ומשומשת להכניס N מפתחות לתוך טבלה בגודל B כך ש-B <= N, אזי תוחלת ההתנגשויות הקשורות במפתח x קטנה מ-1. **Perfect Hash**: בניית הטבלה בתוחלת O(n), פעולת Find עולה O(1) WC.

1 = O(n^{1/\log n}) = o(\log(\log n)) = o(\sqrt{\log n}) = o(\log n) = o(\log^2 n) = o(n^{1/3})

= o(n) = o(2^{\log n}) = o(n^2) = o(\log n)! = o(\log n)^{\log n} = o(n^{\log(\log n)}) =

o(3/2^n) = o(2^n) = o(n*2^n) = o(n!)

log(n!) = O(n*logn) n log n = o(n^{1.5}) log_b n ∈ O(lgn)

∑_{i=1}^n 1/i^2 = O(log(log n)) ∑_{i=0}^n x^i = 1/(1-x) ∑_{i=1}^n √i = O(n * √n) ∑_{i=1}^n i^4 = O(n^{k+1})

∑_{i=1}^n 1/i = O(log n) ∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)

∑_{i=1}^n log i = O(n log n)