

פתרון לתרגיל הראשון

חוזר חלילה

האלגוריתם יעבוד באופן הבא: ראשית נבחר קבוצה $Q \subset \{1, \dots, n\}$ באופן יוניפורמי מכל תתי-קבוצה מגודל q (אחר-כך נבחר את q). לאחר זאת נבחן את הצמצום של f ל- Q (נשאל את $f(i)$ לכל $i \in Q$). נדחה את f אם לכל $1 \leq k \leq n/2$ קיימים $i, j \in Q$ שעבורם $j - i$ מתחלק ב- k אבל $f(i) \neq f(j)$, ואחרת נקבל את f .

ברור שכל פונקציה מחזורית תתקבל ע"י האלגוריתם בהסתברות 1 (לא יהיו קיימים שום i, j כמו למעלה עבור k שהוא המחזור של הפונקציה). נותר להוכיח, עבור בחירה מתאימה של $q = O(\sqrt{n \log(n)/\epsilon})$, שפונקציה ϵ -רחוקה תידחה בהסתברות לפחות $\frac{1}{2}$.

יהיה יותר נוח להשתמש בשאלה "לבלוע את החוכמה" ולנתח את המצב עבור קבוצה Q' שמוגדרת ע"י כך שכל $i \in \{1, \dots, n\}$ יכנס ל- Q' בהסתברות p באופן ב"ת. אנחנו נראה עבור p מתאים שהסתברות הדחיה היא לפחות $\frac{2}{3}$, ואז אפשר לבחור $q = \lceil 6np \rceil$ ולהשתמש בתוצאה של "לבלוע את החוכמה" עבור Q .

נשים לב שאפשר להפוך כל פונקציה $f : \{1, \dots, n\} \rightarrow \{0, 1\}$ למחזורית עם מחזור k באופן הבא: לכל $1 \leq i \leq k$, נסתכל על הקבוצה $J_i = \{i, i+k, i+2k, \dots, i+k\lfloor \frac{n-i}{k} \rfloor\}$. אנחנו נבדוק מהו הערך הכי נפוץ בצמצום $f|_{J_i}$ (אם שני הערכים האפשריים מופיעים אותו מספר של פעמים אז בוחרים שרירותית באחד מהם), ונשנה את כל ערכי f על J_i להיות זהים לו. אם נסמן ב- l_i את מספר הערכים על J_i השונים מהערך הכי נפוץ, נקבל מכך שעבור פונקציה f שהיא ϵ -רחוקה ממחזוריות (ולכן בפרט ϵ -רחוקה מלהיות פונקציה עם מחזור k) מתקיים $\sum_{i=1}^k l_i \geq \epsilon n$.

על מנת לחסום את הסיכוי לגלות את החריגה, לכל $1 \leq i \leq k$ נקבע זיווג (לא מושלם) בגודל l_i בין כל האיברים ב- J_i עם הערך הפחות נפוץ לבין האיברים עם הערך הנפוץ (לא בהכרח כולם). סה"כ בידנו $\sum_{i=1}^k l_i \geq \epsilon n$ זוגות זרים. הסיכוי לא לגלות שני ערכים שונים בתוך אותו J_i אינו עולה על הסיכוי לא להכיל אף אחד מהזוגות הנ"ל, אשר חסום ע"י $(1-p^2)^{\epsilon n} \leq e^{-\epsilon p^2 n}$.

עבור בחירה של $p = 2\sqrt{\log(n)/\epsilon n}$ למשל, הביטוי הנ"ל קטן מ- $1/3n$. באמצעות החסם על איחוד מאורעות עבור $\lfloor n/2 \rfloor$ ערכי k האפשריים, מקבלים שפונקציה ϵ -רחוקה ממחזוריות עבור כל k אפשרי תידחה בהסתברות לפחות $\frac{2}{3}$, כנדרש. לבסוף נחזור ונציב (עבור המעבר לבחירה יוניפורמית של קבוצת שאילתות מגודל קבוע) $q = \lceil 6np \rceil = \lceil 12\sqrt{n \log(n)/\epsilon} \rceil$ להשלמת ההוכחה.

לשיר סולו

ננתח את האלגוריתם הבא: ראשית נגדיל חלוקה של $\{1, \dots, n\}$ לשתי קבוצות זרות A, B , ע"י כך שכל $1 \leq i \leq n$ יוגרל, באופן ב"ת באינדקסים אחרים, להיות ב- A בהסתברות $\frac{1}{2}$ ולחיות ב- B בהסתברות $\frac{1}{2}$.

בשלב הבא נגדיל שני וקטורים $z, y \in \{0, 1\}$ באופן הבא: עבור z , לכל $i \in A$ נגדיל את z_i באופן יוניפורמי וב"ת מתוך $\{0, 1\}$, ולכל $i \in B$ נקבע $z_i = 0$. עבור y , באופן ב"ת בהגרלות עבור z , לכל $i \in B$ נגדיל את y_i באופן יוניפורמי וב"ת מתוך $\{0, 1\}$, ולכל $i \in A$ נקבע $y_i = 0$.

לבסוף, אם $f(z) = f(y) = 1$ אז נדחה את f , ובכל מקרה אחר נקבל אותו.

נטען עתה שהאלגוריתם יקבל פונקציה שהיא פונקציה האפס או פונקציה מהצורה $f(x) = x_i$ בהסתברות 1, וידחה כל פונקציה לינארית אחרת בהסתברות לפחות $\frac{1}{8}$. עבור השלמת השאלה אפשר עתה להריץ את האלגוריתם הנ"ל עשר פעמים (עם הגרלות ב"ת) ולדחות את f אם הוא דחה בלפחות אחת מהן.

הסיבה שפונקציה האפס תתקבל תמיד היא ברורה, וגם פונקציה מהצורה $f(x) = x_i$ תתקבל תמיד, כי לא יתכן שיתקיים גם $i \in A$ וגם $i \in B$. עבור פונקציה לינארית התלויה במספר יותר גדול של משתנים, הסיבה שהיא תידחה בהסתברות לפחות $\frac{1}{8}$ נובעת מהניתוח הבא.

- פונקציה לינארית שאינה מהצורה למעלה חייבת להיות תלויה בלפחות שני משתנים. נסמן שניים מהם ב- x_i ו- x_j . בהסתברות $\frac{1}{2}$, אחד מ- i ו- j נבחר להיות ב- A בעוד שהשני מהם נבחר להיות ב- B .
 - נניח שהמאורע מהסעיף הראשון התקיים, ובלי הגבלת הכלליות נניח שמתקיים $i \in A$ ו- $j \in B$. בהסתברות $\frac{1}{2}$ יתקיים $f(z) = 1$. על מנת לראות זאת, עבור z שהוגרל נסמן את z' שמתקבל ע"י הפיכה של z_i (ז"א ש- $z'_i = 1$ אם ורק אם $z_i = 0$). הוקטור z' מתפלג בדיוק כמו z , ולכן $\Pr[f(z) = 1] = \Pr[f(z') = 1]$. מצד שני, בגלל שהמדובר בפונקציה לינארית מעל $\mathbb{Z}_2$, בדיוק אחד מ- $f(z)$ ו- $f(z')$ יהיה שווה ל-1. לכן ההסתברות עבור $f(z) = 1$ שווה ל- $\frac{1}{2}$.
 - שוב בהנחה שהמאורע מהסעיף הראשון התקיים, ניתוח זה מראה שבהסתברות $\frac{1}{2}$ יתקיים $f(y) = 1$. בנוסף לכך, המאורעות $f(z) = 1$ ו- $f(y) = 1$ הם ב"ת זה בזה בהתניה על קיום המאורע מהסעיף הראשון.
- מאלו נובע שההסתברות שיתקיים $f(z) = f(y) = 1$ והקלט ידחה היא לפחות מכפלת ההסתברויות הנ"ל, $\frac{1}{8}$.

פתרון לתרגיל השני

לרקוד סולו

נבדוק את הפתרון הרשמי של השאלה "לשיר סולו" מהתרגיל הקודם, ונשפר את הסתברות ההצלחה שלו מ- $\frac{2}{3}$ ל- $\frac{5}{6}$ (אפשר למשל לבצע את הפרוצדורה המרכזית שמופיעה שם 15 במקום 10 פעמים). נסמן עתה את מספר השאילות של האלגוריתם הזה (שעובד רק עבור פונקציות לינאריות) ב- q .

מפתה עתה להשתמש באלגוריתם דו-שלבי: קודם כל לבצע בדיקה עבור התכונה שהפונקציה f היא לינארית, ולאחריו את האלגוריתם של "לשיר סולו". הבעיה היא שהתפלגות השאילות של האלגוריתם של "לשיר סולו" אינה יוניפורמית (אפילו כל שאילתה לכשעצמה), ולכן אפילו בפונקציה קרובה ללינארית יכול להיות שהוא יטעה בהסברות גבוהה.

נעשה אם כן שינוי נוסף: בכל פעם שהאלגוריתם של "לשיר סולו" מנסה לבצע שאילתה " $f(x)$ ", במקום זאת נגריל ערך $w \in \{0, 1\}^n$ באופן יוניפורמי וב"ת תלוי שאר השלבים (כולל בשאילות אחרות), ונחשב את הערך של "הפונקציה המתוקנת" $f_w(x)$ כפי שהיא מופיעה בהרצאה על בדיקת הלינאריות. החישוב הזה מבצע שתי שאילות שכל אחת מהן לחד מתפלגת באופן יוניפורמי, ולכן אם f היא ϵ -קרובה לפונקציה לינארית g , בהסתברות לפחות $1 - 2\epsilon$ יתקבל $f_w(x) = g(x)$. האלגוריתם הסופי יהיה האלגוריתם הבא:

- ראשית מבצעים $\min\{\epsilon, 1/12q\}$ בדיקה של f עבור התכונה של לינאריות. מספר השאילות של השלב הזה תלוי רק ב- ϵ וב- q (שהוא קבוע גלובלי).

- עתה מבצעים את האלגוריתם של "לשיר סולו" עם שני השינויים שנזכרו למעלה. האלגוריתם מבצע סה"כ $2q$ שאילות, ואם f היא $1/12q$ -קרובה לפונקציה לינארית g , בהסתברות לפחות $\frac{5}{6}$ האלגוריתם יקבל (עבור ההרצאה שלו) ערכים נכונים מתוך g . לכן בהסתברות לפחות $\frac{2}{3}$ האלגוריתם יתן את התשובה הנכונה עבור הפונקציה g הקרובה ל- f .

עבור הניתוח, ראשית נשים לב שאם f היא פונקציה האפס או פונקציה מהצורה $f(x) = x_i$ אז היא בפרט לינארית, בפרט תתקבל בהסתברות 1 בשלב הראשון. כמו כן בשלב השני, מכיוון שבצד $f = g$, גם שלב זה יקבל בהסתברות 1 (כפי שהאלגוריתם של "לשיר סולו" מבטיח).

עתה נשים לב מה קורה אם f היא ϵ -רחוקה מפונקציה האפס או פונקציה מהצורה $f(x) = x_i$. אם היא $\min\{\epsilon, 1/12q\}$ -רחוקה מכל פונקציה לינארית, אז השלב הראשון של האלגוריתם ידחה אותה בהסתברות לפחות $\frac{2}{3}$. אחרת, הפונקציה הלינארית הקרובה g בפרט לא יכולה להיות פונקציה האפס או מהצורה $f(x) = x_i$ (אחרת f לא יכולה להיות ϵ -רחוקה), ולכן השלב השני ידחה אותה בהסתברות לפחות $\frac{2}{3}$.

השילוש

בדיקה עם שגיאה דו-כוונית: האינטואיציה העיקרית כאן היא שעבור גרף שהוא אכן איזומורפי לגרף התלת-צדדי השלם המאוזן (עם n צמתים בכל צד), אם אנחנו יודעים על שלושה צמתים u_1, u_2, u_3 שכל אחד מהם שייך לצד אחר, אז אנחנו יכולים לוודא לכל צומת נוסף z לאיזה צד הוא שייך: הוא שייך לצד ה- i אם ורק אם אין לו קשת ל- u_i . כך אנחנו יכולים לדגום זוגות צמתים, ולבדוק כל פעם האם יש ביניהם קשת (צריכה להיות קשת כזו אם ורק אם הם לא משתייכים לאותו צד).

עבור שלושה צמתים "פוטנציאליים", אפשר לדעת האם הם שלישיה כנדרש פשוט ע"י בדיקה האם הם יוצרים משולש בגרף. מכיוון שהמדובר בגרף תלת-צדדי מאוזן, בדגימה של מספר צמתים תהיה שלישיה כזו בהסתברות גבוהה. עתה נכתוב את האלגוריתם הפורמלי וננתח אותו.

- נבחר באופן מקרי ויוניפורמי קבוצה בת 10 צמתים מהגרף. אם יש בה משולש, נסמן את הצמתים של אחד המשולשים ב- u_1, u_2, u_3 ונעבור לשלב הבא. אחרת נעצור את האלגוריתם ונדחה את הגרף מיידי.

• נדגום באופן יוניפורמי וב"ת (עם חזרות) זוגות צמתים $v_1, w_1, \dots, v_t, w_t$ מהגרף, עבור $t = \lceil 4/\epsilon \rceil$. לכל $1 \leq r \leq t$, נמצא i_r, j_r כך ש- v_r אינו שכן של u_{i_r} ו- w_r אינו שכן של u_{j_r} (אם נמצא צומת שאין לו שני שכנים בדיוק ב- u_1, u_2, u_3 אז נשייך אותו באופן שרירותי לצד של u_1). אנחנו נדחה כאן אם יש $1 \leq r \leq t$ שעבורו $i_r = j_r$ ויש קשת בין v_r ל- w_r , או שעבורו $i_r \neq j_r$ ואין קשת בין v_r ל- w_r .

• נדגום באופן יוניפורמי (ובאופן ב"ת בשלבים הקודמים) סדרת צמתים $v_1, \dots, v_s$ עבור $s = \lceil 100/\epsilon^2 \rceil$. לכל $1 \leq r \leq s$ נמצא i_r כך ש- v_r אינו שכן של i_r (ושוב נשייך שרירותית צמתים שאין להם שני שכנים בדיוק ב- u_1, u_2, u_3). לכל $1 \leq j \leq 3$ נחשב את $s_j = |\{r : i_r = j\}|$. אם לכל j מתקיים $(\frac{1}{3} - \frac{1}{6}\epsilon)s \leq s_j \leq (\frac{1}{3} + \frac{1}{6}\epsilon)s$ אז נקבל את הגרף, ואחרת נדחה אותו.

מספר השאילות בשלב הראשון הוא $\binom{10}{2}$ (קבוע שאפילו לא תלוי ב- ϵ). בשלב השני יש $7t = O(1/\epsilon)$ שאילות (לכל זוג צמתים שואלים על הקשת האפשרית ביניהם, ולכל אחד מהם על שלושת הקשתות האפשריות בינו לבין שלשת הצמתים מהסעיף הקודם). בשלב השלישי מבצעים $3s = O(1/\epsilon^2)$ שאילות, ולכן סה"כ מספר השאילות הוא $O(1/\epsilon^2)$.

עבור הוכחת השלמות, נבדוק את מה קורה עבור גרף איזומרפי לגרף התלת-צדדי השלם והמאוזן. לפי חישוב הכלה והפרדה (יש גם שיטות יותר פשוטות לחסום את זה) ההסתברות של השלב הראשון למצוא u_1, u_2, u_3 כנדרש (ולא לדחות) היא $\frac{5}{6} > \frac{\binom{3n}{10} - 3\binom{2n}{10} + 3\binom{n}{10}}{\binom{3n}{10}}$. אם הצמתים הנ"ל נמצאו, אז בכל השלבים הבאים הצמתים יקוטלגו לפי המיקום הנכון שלהם באחד הצדדים יחד עם ה- u_i המתאים. במקרה זה, השלב השני לא ידחה בכלל. בשלב השלישי, נשים לב שלכל $1 \leq j \leq 3$ יתקיים $E[s_j] = \frac{1}{3}s$, ולפי חסימת סטיות גדולות יתקיים $\Pr[(\frac{1}{3} - \frac{1}{6}\epsilon)s \leq s_j \leq (\frac{1}{3} + \frac{1}{6}\epsilon)s] > 1 - 2e^{-200/36} > \frac{99}{100}$. בפרט, בהסתברות העולה על $\frac{5}{6}$ השלב השלישי יקבל.

עתה עבור הוכחת הנאותות נראה מה יקרה עבור גרפים ϵ -רחוקים מאיזומורפיזם לגרף הנ"ל. שימו לב שהמדובר בצורך לשנות לפחות $9\epsilon n^2$ קשתות (כי הגרף הוא עם $3n$ צמתים). אם השלב הראשון ממילא דחה אז אין בעיה. אחרת, לאחר השלב הראשון יש לנו שיוך (זה לא משנה שהוא שרירותי עבור צמתים מסויימים) של כל צומת לאחד משלושת הצדדים לפי ה- u_i המתאים. נסמן לכל צומת $v \in V$ את השיוך שלו ב- i_v , ונסמן גם $n_j = \{v \in V : i_v = j\}$ לכל $1 \leq j \leq 3$.

אם יש לפחות $\frac{9}{2}\epsilon n^2$ זוגות מפרים (זוגות צמתים v, w שעבורם $i_v \neq i_w$ ואין ביניהם קשת, או שעבורם $i_v = i_w$ ויש ביניהם קשת) אז השלב השני ידחה בהסתברות לפחות $\frac{2}{3} > 1 - (1 - \frac{1}{2}\epsilon)^{4/\epsilon}$. אחרת, אפשר לשנות את הגרף בלא יותר ב- $\frac{9}{2}\epsilon n^2$ מקומות ולהפוך אותו לגרף תלת-צדדי מלא (אם כי לא בהכרח מאוזן).

בנוסף, אם יש j שעבורו $n_j > (1 + \epsilon)n$ אז יתקיים בשלב השלישי $E[s_j] > \frac{1}{3}(1 + \epsilon)s$ ואז שלב זה ידחה בהסתברות לפחות $\frac{2}{3}$. אותו דבר יקרה אם יש j שעבורו $n_j < (1 - \epsilon)n$. אם אף אחד מאלו לא מתקיים, אז "א שלכל j מתקיים $(1 - \epsilon)n \leq n_j \leq (1 + \epsilon)n$ ", אז אפשר לשנות את הגרף התלת-צדדי המלא מסוף הדיון בשלב השני בלא יותר מ- $\frac{9}{2}\epsilon n^2$ מקומות נוספים ולהפוך אותו למאוזן, מה שמהווה סתירה להנחה שהגרף שלנו היה רחוק. הסיבה לכך היא שאנחנו יכול "להזיז" לא יותר מ- $\frac{3}{2}\epsilon n$ צמתים (תוך שינוי אפשרי של כל הקשתות שלהם) בין הצדדים של הגרף ולהפוך אותו למאוזן. זאת כי מתקיים $\sum_{j=1}^3 n_j = 3n$ ולכן מתקיים $\sum_{\{j: n_j > n\}} (n_j - n) = \sum_{\{j: n_j < n\}} (n - n_j) = \frac{1}{2} \sum_{j=1}^3 |n - n_j|$.

פתרון לתרגיל השלישי

שילוש באחד

עבור החסם התחתון, ראשית נשים לב שאם האלגוריתם מקבל פרמוטציה שנבחרה מקרית ויוניפורמית של הגרף התלת-צדדי המלא המאוזן, אז כל עוד האלגוריתם ביצע פחות מ- $\frac{1}{2}n$ שאילתות, יש סיכוי (שיכול להיות קטן מאוד) שכל השאילתות יחזירו "אין קשת". לכן, אלגוריתם בדיקה עם שגיאה חד-כיוונית שמבצע כזה מספר של שאילתות תמיד חייב לקבל את הגרף אם אף שאילתה לא החזירה שיש קשת. מצד שני, אם האלגוריתם יקבל כקלט את הגרף חסר הקשתות (שהוא בפרט $\frac{1}{2}$ -רחוק מלהיות תלת-צדדי מלא ומאוזן), אז בהסתברות 1 כל השאילתות יחזירו "אין קלט", ולכן הוא יהיה חייב לקבל גם את הגרף הזה. עבור החסם העליון, נבצע את האלגוריתם הבא:

- ראשית ניקח צומת שרירותי u_1 ונשאל על הקשתות ממנו לכל הצמתים האחרים בגרף (סה"כ $3n - 1$ שאילתות). אם אין לו שכן אז נעצור ונדחה את הגרף מיידית, אחרת נסמן את אחד השכנים ב- u_2 .
- עתה נשאל על כל הקשתות מ- u_2 לצמתי הגרף האחרים (עוד $3n - 2$ שאילתות, כבר עשינו בשלב הקודם את השאילתה (u_1, u_2)). אם אין לו שכן שהוא גם שכן של u_1 אז נדחה את הגרף מיידית, ואחרת נסמן את אחד הצמתים הנ"ל ב- u_3 .
- עתה נשאל על כל הקשתות מ- u_3 לצמתי הגרף האחרים (עוד $3n - 3$ שאילתות). אם יש איזה שהוא צומת בגרף שאינו מחובר לבדיוק שניים משלושת הצמתים u_1, u_2, u_3 , נדחה מיידית. כמו כן, נדחה מיידית אם יש $1 \leq i \leq 3$ שאינו מחובר ל- $2n$ צמתים אחרים בדיוק.
- לבסוף, נבצע את השלב השני של האלגוריתם המקורי לשאלה "השילוש", עם u_1, u_2, u_3 שסומנו כאן. נדחה אם השלב הזה דחה, ואחרת נקבל. שימו לב שזהו השלב ההסתברותי היחיד באלגוריתם.

מספר השאילתות הכולל כאן הוא $O(n + 1/\epsilon)$. אם הגרף הוא אכן תלת-צדדי מלא ומאוזן, אז בפרט שני השלבים הראשונים לא ידחו וימצאו u_1, u_2, u_3 שכל אחד מהם נמצא בצד אחר, והשלב השני גם לא ידחה (כי צומת שנמצא בתד של u_i יהיה מחובר לשני הצמתים האחרים ולא ל- u_i). השלב האחרון מתייחס לשלב באלגוריתם של "השילוש" שלפי הניתוח שם תמיד יקבל כאשר u_1, u_2, u_3 מתייחסים לשלושה צדדים של גרף תלת-צדדי מלא (אפילו אם הוא לא מאוזן).

נניח עתה שהגרף הוא ϵ -רחוק מתלת צדדי שלם ומאוזן. אם האלגוריתם הגיע לשלב השלישי וזה קיבל, אז נתייחס לסימונים i_v לכל $v \in V$ ו- $1 \leq j \leq 3$ כפי שאלו הוגדרו בפתרון הרשמי של "השילוש". קבלה ע"י השלב השלישי בהכרח אומרת שמתקיים $n_1 = n_2 = n_3 = n$. על כן אם הגרף הוא ϵ -רחוק אז קיימים לפחות $3\epsilon n^2$ זוגות v, w שמפרים את החלוקה הנ"ל (ז"א שעבורם או $i_v = u_w$ ויש ביניהם קשת, או $i_v \neq i_w$ ואין ביניהם קשת). אבל אז השלב האחרון ידחה בהסתברות גבוהה מ- $\frac{2}{3}$.

אדפטיביות בסדר

כזכור אפשר לראות כל אלגוריתם בדיקה הסתברותי בעל q שאילתות כמרחב הסתברות מעל אלגוריתמים דטרמיניסטים בעלי q שאילתות. לכן מספיק להראות איך אפשר לתרגם אלגוריתם דטרמיניסטי מבוסס-סדר בעל q שאילתות לאלגוריתם לא-אדפטיבי בעל $(1 + o(1)q)!$ שאילתות.

נעשה זאת באופן דומה להדגמה של "פרישת עץ ההחלטה" מהתרגול, אבל כאן יהיה חסם אחר לגודל של עץ ההחלטה עצמו. גם כאן שורש העץ יכול את מיקום השאילתה הראשונה שהאלגוריתם ביצע (לפני שהיה לו מידע כל שהוא לבסס עליו את ההחלטה). לשורש יהיה אבל בן יחיד, כי לפני השאילתה השנייה עדיין לא היתה לאלגוריתם אף שאילתה שניתן להשוות עם השאילתה הראשונה.

באופן כללי, נראה כמה ילדים יכולים להיות לצומת v ברמה ה- i בעץ ההחלטות, ז"א לאחר ביצוע i שאילתות. מסלול ההגעה ל- v (בעל $i - 1$ קשתות) כבר מתאר את כל ההשוואות בין $i - 1$ השאילתות הקודמות (אפשר

להניח שעם כל שאילתה קיבלנו מיידית את כל ההשוואות בינה לבין השאילתות הקודמות וכבר פעלנו לפי המידע הזה). הערך של f על השאילתה של v עתה מושווה מול ערכים של $i-1$ שאילתות קודמות. נניח שאנחנו ממיינים את הערכים הנ"ל לפי $w_1 \leq w_2 \leq \dots \leq w_{i-1}$ (האלגוריתם עצמו לא יודע את הערך w_j , אבל הוא יודע איזו מהשאילתות שלו נתנה את w_j , וכן אלו מאי-השוויונות הנ"ל הם שוויונות). הערך החדש יכול להיות קטן מ- w_1 (ואז קטן מכל הערכים), גדול מ- w_{i-1} (ואז גדול מכולם, שווה לאחד מ- w_j עבור $1 \leq j \leq i-1$ כל שהוא, או גדול מ- w_j וקטן מ- w_{j+1} עבור $1 \leq j \leq i-2$ כל שהוא. סה"כ המדובר הוא לכל היותר ב- $2i-1$ אפשרויות (פחות אפשרויות אם יש שוויונות בערכים קודמים), ולכן לצומת ברמה ה- i יש לכל היותר מספר כזה של בנים.

על כן המספר הכולל של צמתים נושאי שאילתות (צמתים שאינם עלים) בעץ ההחלטה של האלגוריתם חסום ע"י $\sum_{i=1}^q \prod_{j=1}^{i-1} (2j-1) \leq q \prod_{j=1}^{q-1} (2j-1) \leq 2^q \cdot q!$. על מנת להפוך את האלגוריתם לאלגוריתם לא-אדפטיבי, אנחנו נבצע את כל השאילתות שמופיעות בכל צמתי עץ ההחלטה מראש, ורק אז נעקוב אחרי המסלול ה"נכון" מהשורש לעלה המתאים, בדיוק כפי שהודגם בתרגול. על מנת לקבל את הביטוי " $((1+o(1)q)!$ ", נשים לב שלכל $\alpha > 1$ ולכל $q > \max\{\frac{1}{\alpha-1}, 2^{1/(\alpha-1)}\}$ מתקיים $q! \geq 2^q \cdot q! \geq q^{(\alpha-1)q} \cdot q! \geq ([\alpha q])!$.

פתרון לתרגיל הרביעי

שולים מוקש

ראשית, נטען שאם גרף G בעל n צמתים ודרגות חסומות ע"י d הוא ϵ -רחוק (במודל הדליל) מלא להכיל את H כתת-גרף, אז הוא מכיל לפחות $C \cdot \epsilon n$ עותקים זרי-צמתים של H עבור C שתלוי רק ב- m ו- d . אם יש לא יותר מ- k עותקים זרי-צמתים של H , נבחר משפחה מקסימלית של קבוצות צמתים זרות $U_1, \dots, U_l$ (עבור $l \leq k$) של קבוצות צמתים מגודל m כך שעל כל U_i יש עותק של H (הקבוצה מקסימלית בכך שכל עותק אחר של H בהכרח יכיל לפחות צומת אחד ששייך לאחת הקבוצות הנ"ל). עתה נסתכל על הגרף G' שמתקבל מ- G ע"י הסרת כל הקשתות שנוגעות בכל צומת שהוא מתוך האיחוד $\bigcup_{i=1}^l U_i$. סה"כ מספר הקשתות שהסרנו אינו עולה על $d \cdot \big|\bigcup_{i=1}^l U_i\big| \leq dkm$, ובגרף G' לא יכול להיות עותק של H (הסרת קשתות לא יוצרת עותקים חדשים של H , וב- G' לא יכול להיות עותק שמכיל צומת מ- $\bigcup_{i=1}^l U_i$ כי אלו הפכו לצמתים מנותקים, בעוד H קשיר ולכן לא מכיל צמתים כאלה). על כן, אם G הוא ϵ -רחוק מלא-להכיל עותק של H , אז חייב להתקיים $dkm \geq \epsilon dn$, ובפרט הטענה נכונה עבור $C = 1/km$.

על כן, בגרף ϵ -רחוק חייבים להיות לפחות $Cm \cdot \epsilon n$ צמתים ששייכים לעותקים של H (אלו השייכים ללפחות $C \cdot \epsilon n$ עותקים זרי-צמתים). אלגוריתם בדיקה יעבוד בצורה הבאה: בוחרים $\lceil 2/\epsilon Cm \rceil$ צמתים של G באופן יוניפורמי וב"ת. לכל צומת, מבצעים חיפוש לרוחב באורך m ובודקים האם הקבוצה המתקבלת מכילה עותק של H . אם לא מצאנו עותק כזה אז מקבלים, ואחרת דוחים. מספר הצמתים בכל חיפוש חסום ע"י $\sum_{i=0}^m d^i < d^{m+1}$, ולכן מספר השאילתות בחיפוש חסום ע"י d^{m+2} ומספר השאילתות הכולל חסום ע"י $\lceil 2/\epsilon Cm \rceil d^{m+2}$, סכום שאינו תלוי ב- n . ברור מדוע האלגוריתם תמיד יקבל גרף ללא עותק של H (דחיה מתבצעת רק אם עותק כזה נמצא). אם לעומת זאת הגרף G הוא ϵ -רחוק מהתכונה הזו, אז הסיכוי לא לתפוס צומת ששייך לעותק של H חסום ע"י $(1 - Cm \cdot \epsilon)^{\lceil 2/\epsilon Cm \rceil} < \frac{1}{3}$, וברגע שביצענו חיפוש מצומת ששייך לעותק של H אנחנו נמצא את העותק ונדחה את הגרף (כאן השתמשנו בנתון ש- H קשיר).

שולים מוקשים

ראשית נראה פתרון עבור $k = 1$, ז"א כאשר צריך לבדוק את הקלט G עבור אי-הכלה של גרף אחד H . אם H קשיר אז השאלה זהה לשאלה הקודמת, ולכן נתמקד במקרה ש- H אינו קשיר. נסמן את רכיבי הקשירות שלו ב- $L_1, \dots, L_m$. אלגוריתם הבדיקה יעבוד ב- m שלבים באופן הבא: בשלב ה- i , האלגוריתם יבצע $\epsilon/2$ -בדיקה עבור אי הכלה של L_i , לפי הפרוצדורה של השאלה הקודמת, אבל עם הסתברות שגיאה $\frac{1}{3m}$ במקום $\frac{1}{3}$ (אפשר למשל להשיג את זה ע"י $\lceil \log m / \log 3 \rceil + 1$ הרצות ב"ת של האלגוריתם עם הסתברות שגיאה $\frac{1}{3}$). אם לא נמצא עותק כזה, אז מקבלים את הגרף ועוצרים מיידית. אחרת, מסירים מהגרף את הצמתים של העותק של L_i שנמצא ואז עוברים לשלב ה- $i+1$. אם בכל אחד מהשלבים מצאנו עותק של ה- L_i המתאים, נדחה את הגרף.

האלגוריתם לעולם לא ידחה גרף שאין בו עותק של H , כי דחיה פירושה שיש עותקים זרי-צמתים של $L_1, \dots, L_m$ ולכן של H (כזכור בסוף כל שלב האלגוריתם גורע את הצמתים של העותק של L_i שנמצא, מה שמבטיח שהעותקים הנ"ל זרי-צמתים).

נניח שהאלגוריתם מקבל את הגרף בהסתברות לפחות $\frac{1}{3}$. זה אומר שיש i כך שבהסתברות לפחות $\frac{1}{3m}$ האלגוריתם מגיע לשלב ה- i ואז מקבל. על כן יש גם עותקים ספציפים של $L_1, \dots, L_{i-1}$ שאם מסירים אותם מהגרף, אז הגרף הנוצר עובר את הבדיקה נגד L_i בהסתברות לפחות $\frac{1}{3m}$, ולכן אינו $\epsilon/2$ -רחוק מלהיות חסר עותק של L_i . אפשר עתה להפוך את G לגרף חסר עותק של H באופן הבא: את כל הצמתים המשתתפים בעותקים שמצאנו של $L_1, \dots, L_{i-1}$ נבודד (נסיר את כל קשתותיהם), ובשארית הגרף (שיש בה לא יותר מ- n צמתים) נבצע לא יותר מ- $\epsilon dn/2$ שינויים להפוך אותה לחסרת L_i . זה אומר שגם בגרף הכולל שהתקבל אין עותק של L_i (כאן השתמשנו בהנחה ש- L_i אינו צומת מבודד כי אין כאלה ב- H), ולכן גם אין בו עותק של H . אם r מספר הצמתים הכולל של H , אז מספר השינויים הכולל הוא לכל היותר $\epsilon dn/2 + rd$, ולכן אם מספר הצמתים של G הוא לפחות $2r/\epsilon$ אז המדובר בלא-יותר מ- ϵdn שינויים סה"כ.

עתה ננסח אלגוריתם בדיקה עבור אי-הכלה של אף גרף מתוך משפחה $H_1, \dots, H_k$. אנחנו נעבור על הגרפים לפי סדרם, ולכל $1 \leq j \leq k$ נבצע ϵ/k -בדיקה עבור אי-הכלה של H_j , עם הסתברות שגיאה לפחות $\frac{1}{3}$. אנחנו נדחה אם לפחות אחד מהבדיקות הנ"ל דחתה, ואחרת נקבל.

אם הגרף לא מכיל עותקים של אף אחד מ- $H_1, \dots, H_k$, אז כל הבדיקות יקבלו, ולכן הגרף יתקבל. עתה נניח שהאלגוריתם קיבל את הגרף G בהסתברות לפחות $\frac{1}{3}$, ונוכיח שזה אומר שהגרף הוא ϵ -קרוב לאי הכלה של אף אחד מ- $H_1, \dots, H_k$. המשמעות של הסתברות הקבלה היא שכל אחד ואחד מהבדיקות מקבלת בסתברות לפחות $\frac{1}{3}$, ולכן הגרף G הוא ϵ/k -קרוב לאי-הכלת H_j לכל $1 \leq j \leq k$. עתה נשתמש בעובדה שתכונת אי-הכלה של גרף H_j נתון היא מונוטונית, ז"א שהגרף הכי קרוב ל- G שמקיים אותה מתקבל מ- G ע"י הסרה בלבד (בלי הוספה) של לכל היותר $\epsilon dn/k$ קשתות. לכן נוכל לכל $1 \leq j \leq k$ לסמן ב- F_j את קבוצת הקשתות שצריך להסיר על מנת להפוך את G לגרף שאינו מכיל את H_j , ואז להסיר מ- G את איחוד הקבוצות $\bigcup_{j=1}^k F_j$. באופן זה הסרנו לא יותר מ- ϵdn קשתות מ- G סה"כ, והגרף המתקבל G' אינו מכיל אף אחד מהגרפים $H_1, \dots, H_k$, כי לכל $1 \leq j \leq k$ הגרף G' הוא תת-גרף של גרף שאינו מכיל את H_j .

עולים על מוקש

הדוגמה המתאימה היא משפחת כל המעגלים מגודל אי-זוגי (ז"א $H_i = C_{2i+1}$ לכל $i \leq 1$). אי הכלת אף גרף מהמשפחה הזו שקולה לדו-צביעות, ולפי מה שלמדנו בתרגול, $\frac{1}{300}$ -בדיקה לתכונה זו דורשת $\Omega(\sqrt{n})$ שאילתו במודל הדליל עם $d = 3$.

לינאריות באמת

הבדיקה עבור תכונת הלינאריות תהיה עם שלושה שלבים:

- ראשית, נבדוק שמתקיים $f(0) = 0$.
- שנית, נבצע בדיקה ל"כפלויות בסקלר": נבחר $u \in \mathbb{F}^n \setminus \{0\}$ וכן $\alpha \in \mathbb{F} \setminus \{0\}$ באופן יוניפורמי (מהאפשרויות) וב"ת, ונבדוק האם מתקיים $f(\alpha \cdot u) = \alpha \cdot f(u)$.
- לבסוף, נבצע בדיקת חיבוריות כפי שנלמדה בשיעור: נבחר $u, v \in \mathbb{F}^n$ באופן יוניפורמי וב"ת, ונבדוק האם מתקיים $f(u + v) = f(u) + f(v)$.
- אנחנו נדחה את f אם לפחות אחד מהמבחנים למעלה נכשל, ואחרת נקבל.

ברור שאם f היא פונקציה לינארית, אז היא תעבור את כל שלושת המבחנים בהסתברות 1. עתה ננתח, עבור $\epsilon < \frac{5}{8}$ כלשהוא, את הסיכוי לדחות פונקציה f שהיא ϵ -רחוק מלהיות פונקציה לינארית מ- $\mathbb{F}^n$ ל- $\mathbb{F}$. נניח שמתקיים $f(0) = 0$, כי אחרת היא נדחית תמיד (לפי הבדיקה הראשונה) ממילא.

עתה ננתח קירבה לפונקציה שמקיימת כפלויות בסקלר (הבדיקה השניה). נשים לב שאם נגדיר יחס $\sim$ מעל הקבוצה $\mathbb{F} \setminus \{0\}$ לפי שנקבע $u \sim v$ אם ורק אם קיים $\alpha \in \mathbb{F} \setminus \{0\}$ שעבורו $u = \alpha \cdot v$, אז המדובר יהיה ביחס שקילות (סימטריה מתקבלת לפי $v = \alpha^{-1} \cdot u$, רפלקסיביות לפי $u = 1 \cdot u$, ועבור טרנזיטיביות אם $u = \alpha \cdot v$ ו- $v = \beta \cdot w$ אז $u = (\alpha\beta) \cdot w$). כמו כן, כל מחלקת שקילות של $\sim$ מכילה בדיוק אותו מספר איברים, $|\mathbb{F}| - 1$.

נבנה פונקציה g מ- f בצורה הבאה: לכל מחלקת שקילות $A \subset \mathbb{F}^n$ של $\sim$, נבחר $u \in A$ שעבורו יש מספר מקסימלי של $\alpha \in \mathbb{F} \setminus \{0\}$ שמקיימים $f(\alpha \cdot u) = \alpha \cdot f(u)$ (לרב יהיו כמה u כאלו, אז נבחר אחד מהם שרירותית), ולכל $\alpha \in \mathbb{F} \setminus \{0\}$ נגדיר את $g(\alpha \cdot u) = \alpha \cdot f(u)$ (בעצם נשנה את f בכל המקומות ש"לא מסכימים עם מה שצפוי לפי u). המרחק בין g ל- f יהיה לכל היותר ההסתברות של השלב השני של האלגוריתם לדחות, מכיוון שהסיכוי לדחות חסום מלמטה ע"י הסיכוי לדחות כאשר ה- u שנבחר היה אחד מה"מקסימלים" יחסית למחלקת השקילות שלו.

אם המרחק של g מ- f הוא לפחות $\frac{\epsilon}{10}$, זהו אם כן חסם תחתון לסיכוי לדחות. על מנת להשלים את ההוכחה, ננתח את הסיכוי לדחיה עבור f שהיא $\frac{\epsilon}{10}$ קרובה ל- g המתאימה, אבל עדיין ϵ -רחוקה מלינאריות. נראה שבמקרה כזה הסיכוי של השלב השלישי של האלגוריתם לדחות הוא לפחות $\frac{\epsilon}{10}$.

נניח בשלילה שזהו לא המצב. במקרה כזה, ראשית נשים לב שלפי איחוד מאורעות בהסתברות לפחות $\frac{3\epsilon}{10}$, שלושת השאלות של השלב השלישי הן ממקומות שבהן f שווה ל- g . לכן הסיכוי לדחות את g (אם היינו מריצים עליה אלגוריתם בדיקת חיבוריות) היה חסום ע"י $\frac{4\epsilon}{10}$, ולכן g היתה $\frac{6\epsilon}{10}$ -קרובה לפונקציה חיבורית h . הפונקציה h הנ"ל היתה מקיימת גם כפליות, ולכן היתה לינארית לחלוטין: עבור $u \in \mathbb{F}^n \setminus \{0\}$ ו- $\alpha \in \mathbb{F} \setminus \{0\}$ קבועים, נגדיל את v באופן יוניפורמי, ונשים לב שההסתברות ש- h תהיה שווה ל- g בכל הנקודות $v, u-v, \alpha \cdot v, \alpha \cdot (u-v)$ היא לפחות $1 - \frac{16\epsilon}{10} > 0$ (כאן השתמשנו בהנחה $\epsilon < \frac{5}{8}$). נבחר v ספציפי שעבורו המאורעות הנ"ל מתקיימים, ואז נקבל

$$\begin{aligned} h(\alpha \cdot u) &= h(\alpha \cdot v) + h(\alpha \cdot (u - v)) = g(\alpha \cdot v) + g(\alpha \cdot (u - v)) \\ &= \alpha \cdot g(v) + \alpha \cdot g(u - v) = \alpha \cdot h(v) + \alpha \cdot h(u - v) = \alpha \cdot h(u) \end{aligned}$$

כמו כן, לפי אי-שוויון המשולש h הנ"ל היא $\frac{7\epsilon}{10}$ -קרובה ל- f , בסתירה להנחות.