שיטות הסתברותיות ואלגוריתמים - חוברת התרגילים

20 ביוני 2025

חוברת זו מכילה תרגילים נבחרים מהשנים הקודמות של הקורס, עם פתרונות.

מרחק בין התפלגויות

קרבה בין התפלגויות

עבור שתי מידות הסתברות μ,ν מעל $[n]=\{1,\ldots,n\}$ נגדיר את המרחק ביניהן (בספרות מרחק זה קרוי Variation Distance) לפי הנוסחה $d(\mu,\nu)=\frac{1}{2}\sum_{i=1}^{n}|\Pr_\mu[i]-\Pr_\nu[i]|$. שימו לב לתכונות הבאות של המרחק:

- תמיד מתקיים $0\le d(\mu,\nu)\le 1$ (לפי $0\le\sum_{i=1}^{n}|\Pr_\mu[i]-\Pr_\nu[i]|\le\sum_{i=1}^{n}\Pr_\mu[i]+\sum_{i=1}^{n}\Pr_\nu[i]=2$).
- מתקיים $d(\mu,\nu)=0$ אם ורק אם μ ו־ν זהים (בגלל שהמרחק הוא סכום של ערכים מוחלטים של ההפרשים).
- מתקיימים סימטריה $d(\mu,\nu)=d(\nu,\mu)$ ואי־שוויון המשולש $d(\mu,\nu)\le d(\mu,\tau)+d(\tau,\nu)$.
- מתקיים $d(\mu,\nu)=1$ אם ורק אם אין j שעבורו גם $\Pr_\mu[j]>0$ וגם $\Pr_\nu[j]>0$ (אם יש j כזה, אז מתקיים $|\Pr_\mu[j]-\Pr_\nu[j]|<\Pr_\mu[j]+\Pr_\nu[j]$ ולכן מתקיים $\sum_{i=1}^{n}|\Pr_\mu[i]-\Pr_\nu[i]|<2$).

א. הראו שלכל מאורע E מתקיים $|\Pr_\mu[E]-\Pr_\nu[E]|\le d(\mu,\nu)$.

ב. הראו שקיים E עבורו מתקיים $|\Pr_\mu[E]-\Pr_\nu[E]|=d(\mu,\nu)$.

התפלגויות מותנות

נניח ש־X ו־Y הם שני משתנים מקריים (לא ב"ת) מעל אותו מרחב הסתברות, אשר מקבלים ערכים ב־$\{1,\ldots,n\}$. נניח שקיים מאורע E כך ש־$\Pr[E]\ge 1-\epsilon$, ושאם E מתקיים אז $X=Y$ (ז"א $\Pr[X=Y|E]=1$). הראו שהמרחק בין ההתפלגות על ערכי X לבין ההתפלגות על ערכי Y הוא לא יותר מ־ϵ, כלומר הראו כי מתקיים $\frac{1}{2}\sum_{i=1}^{n}|\Pr[X=i]-\Pr[Y=i]|\le\epsilon$.

מכפלה של הסתברויות

יהי μ מרחב הסתברות על קבוצת המחרוזות הבינאריות מאורך k, המוגדר כך שלכל i הביט ה־i נבחר להיות 1 בהסתברות α_i באופן בלתי תלוי בבחירת הביטים האחרים. יהי ν מרחב הסתברות על אותה קבוצה, המוגדר באופן זהה, פרט לכך שלכל i הביט ה־i נבחר להיות 1 בהסתברות β_i (באופן בלתי תלוי באחרים). הראו שהמרחק (variation distance) בין μ ו־ν הוא לכל היותר $\sum_{i=1}^{k}|\alpha_i-\beta_i|$.

קו המשווה

נניח ש־μ ו־ν מרחבי הסתברות מעל $\{1,\ldots,n\}$, ו־E מאורע בעל הסתברות חיובית ב־μ ו־ν כך ששתי ההתפלגויות המותנות על E שוות, ז"א לכל i מתקיים $\Pr_\mu[i|E] = \Pr_\nu[i|E]$. הראו שהמרחק $d(\mu,\nu)$ חסום מלמעלה ע"י $\max\{\Pr_\mu[\neg E], \Pr_\nu[\neg E]\}$.

פתרונות לתרגילים על מרחק בין התפלגויות

קרבה בין התפלגויות

א. נסמן ב־$\neg E$ את המאורע המשלים ל־E. מתקיים $\Pr_\mu[\neg E] - \Pr_\nu[\neg E] = \Pr_\nu[E] - \Pr_\mu[E]$, כי $\Pr_\mu[\neg E] + \Pr_\mu[E] = \Pr_\nu[\neg E] + \Pr_\nu[E] = 1$. מכאן מתקבל:

$$\begin{aligned}
|\Pr_\mu[E] - \Pr_\nu[E]| &= \frac{1}{2}|\Pr_\mu[E] - \Pr_\nu[E]| + \frac{1}{2}|\Pr_\mu[\neg E] - \Pr_\nu[\neg E]| \\
&= \frac{1}{2}|\sum_{i\in E}(\Pr_\mu[i] - \Pr_\nu[i])| + \frac{1}{2}|\sum_{i\notin E}(\Pr_\mu[i] - \Pr_\nu[i])| \\
&\le \frac{1}{2}\sum_{i\in E}|\Pr_\mu[i] - \Pr_\nu[i]| + \frac{1}{2}\sum_{i\notin E}|\Pr_\mu[i] - \Pr_\nu[i]| \\
&= \frac{1}{2}\sum_{i\in[n]}|\Pr_\mu[i] - \Pr_\nu[i]| = d(\mu,\nu)
\end{aligned}$$

ב. נגדיר את $E = \{i \in [n] : \Pr_\mu[i] > \Pr_\nu[i]\}$. במקרה זה מתקיימים

$$\begin{aligned}
|\sum_{i\in E}(\Pr_\mu[i] - \Pr_\nu[i])| &= \sum_{i\in E}(\Pr_\mu[i] - \Pr_\nu[i]) &= \sum_{i\in E}|\Pr_\mu[i] - \Pr_\nu[i]| \\
|\sum_{i\notin E}(\Pr_\mu[i] - \Pr_\nu[i])| &= -\sum_{i\notin E}(\Pr_\mu[i] - \Pr_\nu[i]) &= \sum_{i\notin E}|\Pr_\mu[i] - \Pr_\nu[i]|
\end{aligned}$$

ולכן מתקבל שוויון בפיתוח למעלה, ז"א $d(\mu,\nu) = |\Pr_\mu[E] - \Pr_\nu[E]|$.

התפלגויות מותנות

לפני שנמשיך, נשים לב שהטענה מתקיימת מיידית אם $\Pr[E] = 1$ (כי אז $X = Y$ בהסתברות 1, ובפרט הם מתפלגים זהה), וגם אם $\Pr[E] = 0$ (כי אז $\epsilon = 1$, ולכן אנחנו לא טוענים כלום על המרחק בין μ ל־ν).

עבור שאר המקרים משתמשים בנוסחת ההסתברות השלמה, נוסחת ההסתברות המותנה, בשוויון ההסתברויות $\Pr[X = i|E] = \Pr[Y = i|E]$ שנובע מהנתון $\Pr[X = Y|E] = 1$, ולבסוף באי שוויון המשולש (הטיפול המיוחד בשני מקרי הקצה למעלה נעשה כך שכל ההתניות בפיתוח הן על מאורעות עם הסתברות גדולה מ־0):

$$\begin{aligned}
\frac{1}{2}\sum_{i=1}^{n}|\Pr[X=i] - \Pr[Y=i]| &= \frac{1}{2}\sum_{i=1}^{n}\Big|\Pr[(X=i)\wedge E] + \Pr[(X=i)\wedge\neg E] \\
&\qquad -\Pr[(Y=i)\wedge E] - \Pr[(Y=i)\wedge\neg E]\Big|
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{i=1}^{n}\Big|\Pr[X=i|E]\Pr[E]+\Pr[X=i|\neg E]\Pr[\neg E] \\
&\qquad -\Pr[Y=i|E]\Pr[E]-\Pr[Y=i|\neg E]\Pr[\neg E]\Big| \\
&= \frac{1}{2}\sum_{i=1}^{n}\Big|\Pr[X=i|\neg E]-\Pr[Y=i|\neg E]\Big|\Pr[\neg E] \\
&\le \frac{1}{2}\left(\sum_{i=1}^{n}\Pr[X=i|\neg E]+\sum_{i=1}^{n}\Pr[Y=i|\neg E]\right)\Pr[\neg E] \\
&\le \frac{1}{2}\cdot 2\cdot\epsilon=\epsilon
\end{aligned}
$$

מכפלה של הסתברויות

אפשר לפתור את השאלה באמצעות חישוב ישיר, אולם במקום זאת נראה כאן דרך המשתמשת בתוצאות שהוכחנו בתרגילים הקודמים. נגדיר מרחב הסתברות על זוגות של מחרוזות $x_1,\ldots,x_k,y_1,\ldots,y_k$ באופן הבא: לכל i נבחר את x_i ואת y_i באופן ב"ת בבחירות עבור i אחרים (אולם באופן תלוי זה בזה). אם $\alpha_i>\beta_i$ אז בסיכוי β_i נבחר $x_i=y_i=1$, בסיכוי $\alpha_i-\beta_i$ נבחר $x_i=1$ ו־$y_i=0$, ובסיכוי $1-\alpha_i$ נבחר $x_i=y_i=0$. אם $\alpha_i\le\beta_i$ אז בסיכוי α_i נבחר $x_i=y_i=1$, בסיכוי $\beta_i-\alpha_i$ נבחר $y_i=1$ ו־$x_i=0$, ובסיכוי $1-\beta_i$ נבחר $x_i=y_i=0$.

נשים לב עתה שההתפלגות (הלא־מותנה) על $x_1,\ldots,x_k$ זהה ל־μ, ושההתפלגות (הלא־מותנה) על $y_1,\ldots,y_k$ זהה ל־ν. כמו כן, לכל i מתקיים $x_i\neq y_i$ בהסתברות $|\alpha_i-\beta_i|$, ולכן ההסתברות ש־$x_1,\ldots,x_k$ אינה זהה ל־$y_1,\ldots,y_k$ חסומה ע"י $\sum_{i=1}^{k}|\alpha_i-\beta_i|$ לפי איחוד מאורעות. מכאן (ע"פ התרגיל על התפלגויות מותנות) נובע שזהו חסם על המרחק בין μ ל־ν.

הערה: אם נשים לב שהמאורעות $x_i\neq y_i$ ב"ת זה בזה, נוכל לחסום את המרחק בין μ ל־ν על ידי החסם המשופר $1-\prod_{i=1}^{k}(1-|\alpha_i-\beta_i|)\le\sum_{i=1}^{k}|\alpha_i-\beta_i|$ (אי־השוויון יהיה חד אם קיימים לפחות שני ערכי i שעבורם $\alpha_i\neq\beta_i$).

קו המשווה

אפשר לפתור את השאלה ע"י חישוב ישיר, אבל כאן נשתמש בתוצאת השאלה "התפלגויות מותנות" מקודם. הרעיון יהיה להגריל ערכים עבור זוג של משתנים מקריים שייצגו את שתי ההתפלגויות, תוך כדי שדואגים לכך שבהתפלגות המשותפת תהיה הסתברות גבוהה למאורע שהם שווי־ערך.

נגדיר מרחב הסתברות τ מעל $\{1,\ldots,n\}\times\{1,\ldots,n\}$ ומעליו נגדיר שני משתנים מקריים: X שיתפלג כמו μ ו־Y שיתפלג כמו ν, ומאורע מתאים F. נניח בלי הגבלת הכלליות שמתקיים $\Pr_\mu[E]\le\Pr_\nu[E]$ – אחרת קודם נחליף בין μ ל־ν.

- על מנת להגריל את (i,j) לפי τ: בהסתברות $\Pr_\mu[E]$ נגריל k לפי ההתפלגות של μ המותנה על E (שהיא זהה להתפלגות של ν המותנה על E), ונקבע $i=j=k$. בהסתברות $\Pr_\nu[E]-\Pr_\mu[E]$ (אם זה גדול מ־0) נגריל את i לפי ההתפלגות המותנה של μ על $\neg E$ (זה יהיה מוגדר כי במקרה הזה בפרט $\Pr_\mu[\neg E]>0$) ובאופן ב"ת את j לפי ההתפלגות המותנה של ν על E. לבסוף, בהסתברות $1-\Pr_\nu[E]$ (אם היא גדולה מ־0) נגריל את i לפי ההתפלגות המותנה של μ על $\neg E$ ובאופן ב"ת את j לפי ההתפלגות המותנה של ν על $\neg E$ (שתי ההתפלגויות המותנות יהיו מוגדרות במקרה זה).
- המ"מ מוגדרים לפי $X(i,j)=i$ ו־$Y(i,j)=j$.
- המאורע F הוא פשוט זה שמתקיים $i=j$.

נראה עתה שמתקיימים התנאים המבוקשים להפעלת תוצאת השאלה "התפלגויות מותנות".

- לכל $1 \le i \le n$ מתקיים

$$\begin{aligned}\Pr_\tau[X=i] &= \Pr_\mu[E]\Pr_\mu[i|E] + (\Pr_\nu[E]-\Pr_\mu[E])\Pr_\mu[i|\neg E] + (1-\Pr_\nu[E])\Pr_\mu[i|\neg E] \\ &= \Pr_\mu[E]\Pr_\mu[i|E] + \Pr_\mu[\neg E]\Pr_\mu[i|\neg E] = \Pr_\mu[i]\end{aligned}$$

- לכל $1 \le j \le n$ מתקיים

$$\begin{aligned}\Pr_\tau[Y=j] &= \Pr_\mu[E]\Pr_\nu[j|E] + (\Pr_\nu[E]-\Pr_\mu[E])\Pr_\nu[j|E] + (1-\Pr_\nu[E])\Pr_\nu[j|\neg E] \\ &= \Pr_\nu[E]\Pr_\nu[j|E] + \Pr_\nu[\neg E]\Pr_\nu[j|\neg E] = \Pr_\nu[j]\end{aligned}$$

- נובע ישירות מההגדרה שמתקיים $\Pr[X=Y|F]=1$.
- מתקיים $\Pr_\tau[F] \ge \Pr_\mu[E] = \min\{\Pr_\mu[E],\Pr_\nu[E]\}$ כי לפי ההגדרה של τ, בהסתברות $\Pr_\mu[E]$ אנחנו במפורש בחרנו ערך k ואז קבענו $i=j=k$. על כן $\Pr_\tau[\neg F] \le \max\{\Pr_\mu[\neg E],\Pr_\nu[\neg E]\}$.

הערה: בחינה של ההוכחה תראה שמתקבל $d(\mu,\nu) < \max\{\Pr_\mu[\neg E],\Pr_\nu[\neg E]\}$ אם ורק אם קיים $j \notin E$ שיש לו הסתברות חיובית גם לפי ν וגם לפי μ.

בניה וניתוח של מרחבי הסתברות

התפלגויות מותנות – הכיוון השני

נניח ש־$p_1,\ldots,p_n$ ו־$q_1,\ldots,q_n$ הם וקטורי התפלגות (סדרות של מספרים אי־שליליים שסכומן 1) שעבורם מתקיים $\frac{1}{2}\sum_{i=1}^n |p_i-q_i| = \epsilon$. הראו שקיים מרחב הסתברות, ושני מ"מ X ו־Y, כך שלכל i מתקיים $\Pr[X=i]=p_i$ ו־$\Pr[Y=i]=q_i$, וכן המאורע $X=Y$ מתקיים בהסתברות $1-\epsilon$ בדיוק.

התאמה זוגית

נתונים שני מרחבי הסתברות μ_1,μ_2 מעל הקבוצה S ושני מרחבי הסתברות ν_1,ν_2 מעל הקבוצה T. נגדיר את $\tau_1 = \mu_1 \times \nu_1$ להיות מרחב הסתברות מעל קבוצת המכפלה $S \times T$ לפי הנוסחה $\tau_1((a,b)) = \mu_1(a)\cdot\nu_1(b)$, ובאופן דומה נגדיר את $\tau_2 = \mu_2 \times \nu_2$ (קל לוודא שאלה אכן פונקציות הסתברות, ואין צורך להוכיח זאת). הראו שהמרחקים בין ההתפלגויות מקיימים $d(\tau_1,\tau_2) \le d(\mu_1,\mu_2) + d(\nu_1,\nu_2) - d(\mu_1,\mu_2)\cdot d(\nu_1,\nu_2)$.

סימולציה של הסתברות

נתון מספר ממשי $0 < \alpha < 1$. הראו אלגוריתם (עם הוכחה) אשר משתמש אך ורק במ"מ מקריים ב"ת שמקבלים ערך מ־$\{0,1\}$ באופן יוניפורמי ("מטבעות הוגנים"), ופולט "1" בהסתברות α בדיוק ו־"0" בהסתברות $1-\alpha$. תוחלת מספר המטבעות שהאלגוריתם משתמש בהם צריכה להיות חסומה ע"י קבוע שאינו תלוי ב־α.

רמז: אפשר לבצע סימולציה של בחירה יוניפורמית של $0 \le \beta \le 1$, ולעצור את הסימולציה ברגע שהשאלה האם $\alpha < \beta$ או $\beta \ge \alpha$ היא בעלת תשובה וודאית.

התבלבלות

נתונות שתי התפלגויות μ_1 ו־μ_2 מעל $S=\{1,\ldots,n\}$. נגדיר התפלגות חדשה $\hat{\mu}$ עם שני משתנים מקריים X,Y באופן הבא: בהסתברות $\frac{1}{2}$, הערכים המוגרלים של X ו־Y יהיו התוצאות של שתי הגרלות ב"ת לפי μ_1. בהסתברות $\frac{1}{2}$, הערכים המוגרלים של X ו־Y יהיו התוצאות של שתי הגרלות ב"ת לפי μ_2. שימו לב שההתפלגות $\hat{\mu}$ והמשתנים X,Y בעצם מוגדרים מעל הקבוצה $S\times S$. הראו שאם $\mu_1\neq\mu_2$ אז המ"מ X,Y **אינם** בלתי־תלויים זה בזה.

עולה עולה

עבור קבוצה S, לכל $1\le p\le 1$ נגדיר את מרחב ההסתברות μ_p מעל קבוצת החזקה שלה $\mathcal{P}(S)$ באופן הבא: על מנת להגריל את $A\subseteq S$, לכל $a\in S$ נכליל את a ב־A בהסתברות p בדיוק, באופן ב"ת לחלוטין באיברים האחרים. במילים אחרות, $\mu_p(A)=p^{|A|}(1-p)^{|S|-|A|}$.

עבור משפחה של קבוצות $\mathcal{A}\subseteq\mathcal{P}(S)$, נגיד שהיא מונוטונית לא־יורדת אם לכל $A\in\mathcal{A}$ ולכל $A\subseteq B\subseteq S$ מתקיים גם $B\in\mathcal{A}$. הראו שעבור משפחה $\mathcal{A}$ מונוטונית לא־יורדת שאינה שווה ל־$\mathcal{P}(S)$ ואינה ריקה, לכל $0\le p<q\le 1$ מתקיים $\Pr_{\mu_p}[\mathcal{A}]<\Pr_{\mu_q}[\mathcal{A}]$. במילים אחרות, הסיכוי להגריל איבר מ־$\mathcal{A}$ עולה כשעוברים מהגרלה לפי p להגרלה לפי q.

הערות: שימו לב שבמקרה הפרטי שבו S היא קבוצת כל הזוגות מתוך $V=\{1,\ldots,n\}$, המרחב μ_p זהה למרחב הגרפים המקריים $G(n,p)$ שמוגדר בחוברת ההרצאות.

ניחושים מול שקרן

קיים מספר טבעי לא ידוע k שאותו צריך לנחש (אין מראש הגבלה על גודלו). השאלה היחידה שמותר לשאול היא מהסוג "האם המספר שווה ל־a?" כאשר a מספר טבעי כל שהוא. אם $a\neq k$ אז התשובה תמיד תהיה "לא", אבל אם $a=k$ אז רק בסיכוי $\frac{2}{3}$ התשובה תהיה "כן", ובסיכוי $\frac{1}{3}$ התשובה בכל זאת תהיה "לא". כתבו אלגוריתם אשר יצליח למצוא את המספר הנכון (ז"א לקבל תשובה של "כן") לאחר ביצוע מספר ניחושים שתוחלתו היא $O(k)$ לכל k, והוכיחו זאת.

וקטורים בהגרלה

בשאלה זו נדון במרחבים לינאריים מעל השדה $\mathbb{Z}_2=\{0,1\}$ (עם חיבור וכפל מודולו 2). מגרילים באופן מקרי, יוניפורמי, וב"ת סדרת וקטורים $v_1,\ldots,v_n\in(\mathbb{Z}_2)^n$. הראו שבסיכוי חסום מלמטה ע"י קבוע גדול מאפס, הווקטורים האלו יהיו בלתי־תלויים (ובפרט יהוו בסיס של המרחב).

קואליציה רחבה

נתונים קבוצה S, מרחב הסתברות μ מעל S, ופונקצית "ערך" $f:\mathcal{P}(S)\to\mathbb{R}^+\cup\{0\}$ (אין הגבלות על f). נתון גם שעבור כל קבוצה $A\subset S$ מגודל m או פחות, אם מגרילים את $a\in S$ לפי μ, אז בהסתברות לפחות $\frac{1}{2}$ קיימת קבוצה $B\subseteq A\cup\{a\}$ שעבורה מתקיים $f(B)>f(A)$. הראו שאם $X_1,\ldots,X_{10m}$ כולם מוגרלים באופן ב"ת לחלוטין מתוך S לפי μ, אז בהסתברות לפחות $\frac{1}{2}$ קיימת קבוצה $I\subseteq\{1,\ldots,10m\}$ מגודל גדול ממש מ־m שעבורה $f(\{X_i:i\in I\})>0$.

רמז ואזהרה: כדאי להגדיר "מאורע גרוע" לכל $I\subset\{1,\ldots,10m\}$ שעבורו $|I|\le m$, ולהראות שבהסתברות גבוהה מספיק אף אחד מהמאורעות הנ"ל לא מתקיים. האזהרה היא שכדאי להיזהר בפורמליזם, זה לא קשה לכתוב תשובה שרק נראית נכונה במבט ראשון.

בחירה בלי מגוון

עבור k ו־c שלמים חיוביים קבועים, וקבוצה A עם n איברים, נתונה צביעה $f : A \to \{1, \ldots, c\}$. נסמן את משפחת כל תתי הקבוצה בנות k איברים של A בסימון $\binom{A}{k} = \{B \subset A : |B| = k\}$. הראו שקיים מרחב הסתברות μ מעל $\binom{A}{k} \cup \{\emptyset\}$ (חישבו על זה כעל בחירה של k איברים בלי חזרות עם אפשרות ל"כישלון"), שמקיים את המאפיינים הבאים עבור קבוצה K שנבחרת לפי μ:

- לכל $a \in A$ מתקיים $\Pr_\mu[a \in K] \leq k/n$ (הבחירה "מרגישה באופן מקומי" כמו בחירה יוניפורמית של k איברים מ־A).
- בהסתברות 1, כל האיברים של K מקבלים ערכים זהים של f (זה נכון באופן ריק כאשר $K = \emptyset$, אבל אנחנו רוצים באופן כללי התפלגות מעל תתי־קבוצות "מונוכרומטיות").
- לכל $\epsilon > 0$ קיים $N(c, k, \epsilon)$ (תלוי גם ב־c ו־k), כך שאם $|A| = n \geq N$ אז $\Pr_\mu[K = \emptyset] \leq \epsilon$ (עבור c, k קבועים ו־n גדול, הסיכוי לכישלון בבחירה הוא $o(1)$).

מרחב הסתברות בסגנון רמזי

עבור k ו־c שלמים חיוביים קבועים, נתונה צביעה ב־c צבעים של קשתות הגרף הדו־צדדי המלא עם n צמתים בכל צד. באופן פורמלי, נתונות קבוצות הצמתים הזרות U, V שעבורן $|U| = |V| = n$, ונתונה הצביעה $f : U \times V \to \{1, \ldots, c\}$ של כל הזוגות האפשריים. הראו שקיים מרחב הסתברות ν מעל תתי־גרפים מושרים, ז"א מרחב שבוחר $X \subset U$ ו־$Y \subset V$, שמקיים את המאפיינים הבאים:

- תמיד $(X, Y) \in \binom{U}{k} \times \binom{V}{k} \cup (\emptyset, \emptyset)$ (או שבוחרים בדיוק k צמתים מכל קבוצה או ש"נכשלים").
- תמיד כל האיברים של $X \times Y$ מקבלים ערכים זהים של f (תמיד מתקבלים תתי־גרפים "מונוכרומטים", כמו במשפט רמזי).
- לכל $u \in U$ ו־$v \in V$ מתקיים $\Pr_\nu[u \in X \wedge v \in Y] \leq k^2/n^2$ (מבחינת הקשתות של הגרף, הבחירה "מרגישה באופן מקומי" כמו בחירה יוניפורמית של X ו־Y מבין תתי־הקבוצות בנות k איברים).
- לכל $\epsilon > 0$ קיים $M(c, k, \epsilon)$ (תלוי גם ב־c ו־k), כך שאם $n \geq M$ אז $\Pr_\nu[(X, Y) = (\emptyset, \emptyset)] \leq \epsilon$ (עבור c, k קבועים ו־n גדול, הסיכוי לכישלון הוא $o(1)$).

הדרכה: מומלץ להשתמש בתוצאת השאלה "בחירה בלי מגוון" עבור פתרון שאלה זו. בדרך אל הפתרון, חישבו קודם על מרחב הסתברות מעל זוגות (X', Y) שבו מתקיים שעבור כל צומת $u \in X'$, האיברים של $\{u\} \times Y$ מקבלים כולם ערכים זהים של f. חישבו איך מגיעים לזוג כזה, ואיך מגיעים ממנו לזוג (X, Y).

דו־קרב תרנגולים

נתונים שני תרנגולים. בכל סיבוב, כל אחד מהם בוחר, בהסתברות $\frac{1}{2}$ ובאופן ב"ת, האם לנקר את חברו. התהליך מפסיק לאחר שבוצע ניקור כל שהוא, ואז השאלה היא האם תרנגול אחד נשאר בריא או ששני התרנגולים נוקרו בו־זמנית. חשבו את הסיכוי שנשאר תרנגול בריא.

דו־קרב תרנגולים עייפים

שוב נתונים לנו שני תרנגולים. הפעם, בכל סיבוב, כל תרנגול שהוא עדיין בריא וערני בוחר באופן ב"ת אחת משלוש פעולות (כל אחת בהסתברות $\frac{1}{3}$): הוא יכול לא לעשות כלום, ללכת לישון עד סוף התהליך, או לנקר את חברו (זה אפשרי גם אם התרנגול השני ישן). התהליך מסתיים כאשר שני התרנגולים ישנים ו/או מנוקרים. חשבו את הסיכוי שהתהליך יסתיים במצב שבו שני התרנגולים בריאים וישנים בשלווה.

רגע לפני

עבור קבוצה סופית לא ריקה S_0 ומספר ממשי $0 < p < 1$ נבצע את התהליך הבא: בשלב ה־i (כאשר מתחילים מ־$i = 0$), בוחרים את S_{i+1} ע"י כך שכל איבר $a \in S_i$ יבחר להיות ב־S_{i+1} בהסתברות p, באופן ב"ת בכל הבחירות האחרות. הראו שבהסתברות לפחות $\frac{2p}{1+p}$ קיים k שעבורו $|S_k| = 1$.

בלי אורך רוח

נגריל מחרוזת $x = (x_1, \ldots, x_n) \in \{0,1\}^n$ מאורך n ע"י כך שלכל i נגריל את x_i באופן יוניפורמי וב"ת באחרים מ־$\{0,1\}$. יהי X המ"מ שמקבל את ה־m המקסימלי שעבורו x מכילה סדרה של m אפסים רצופים (ז"א שקיים $1 \le i \le n+1-m$ כך ש־$x_j = 0$ לכל $i \le j \le i+1+m$). במקרה הקצה ש־x לא מכילה אפסים כלל נקבע $X = 0$. הראו שמתקיים $\mathrm{E}[X] \le \log n + O(1)$ (הלוגריתם בבסיס 2).

רמז: זה עוזר לנתח את המ"מ $Y = \max\{0, X - \lceil \log n \rceil\}$.

עם אורך רוח

עבור ההגרלה והמ"מ של השאלה "בלי אורך רוח", הוכיחו עתה שמתקיים $\mathrm{E}[X] \ge \log n - O(1)$.

הדרכה: שיטה אחת לעשות את זה היא באמצעות הסתכלות על הגרלה אלטרנטיבית של המחרוזות, ע"י כך שמגרילים את מספר האפסים בין כל שני ביטים שווים ל־1 שנמצאים בה (עם ההתפלגויות המתאימות): כעיקרון לוקחים שרשור של הרבה רצפים כאלה אבל אחר כך "מקצצים" את כל מה שנמצא אחרי n הביטים הראשונים. ניתן קודם כל להוכיח את חסם התוחלת בהינתן המאורע שגם אחרי הקיצוץ כמות מספקת של רצפים נשארו רלוונטים, ואח"כ להרחיב אותו לחסם על התוחלת הלא מותנה. אם אתם משתמשים בשיטה זו, עליכם יהיה להסביר מדוע ההגרלה האלטרנטיבית גם נותנת התפלגות יוניפורמית עבור x. בדרך לזה כדאי להשתמש בטענה ש־x מכילה לפחות $n/4$ אחדות בהסתברות $1 - e^{-\Omega(n)}$ (זה נובע למשל מחסמי סטיות גדולות, אבל בשביל הטענה כאן מספיק גם להשתמש בחסם הבינום $\binom{n}{k} \le \left(\frac{en}{k}\right)^k$, שממנו נובע שההסתברות להכלת פחות מ־$n/4$ אחדות חסומה ע"י $2^{-n}\sum_{k=0}^{\lceil n/4 \rceil - 1}\binom{n}{k} < n2^{-n}(\frac{en}{n/4})^{n/4} < n \cdot (\frac{10}{11})^n = e^{-\Omega(n)}$).

לתפוס את הרכבת

נגדיר סדרה אינסופית של משתנים מקריים ומרחב הסתברות באופן הבא: המ"מ X_0 יוגדר כך שבהסתברות 1 מתקיים $X_0 = 0$. לכל $i > 0$, אחרי שהגרלנו וקבענו כבר את הערך של X_{i-1}, בהסתברות $\frac{1}{3}$ (באופן ב"ת בכל ההגרלות הקודמות) נקבע $X_i = X_{i-1} + 1$ ובהסתברות $\frac{2}{3}$ נקבע $X_i = X_{i-1} - 1$. חשבו את ההסתברות שיהיה i כל שהוא שעבורו $X_i = 1$. **נתון לכם שההסתברות הזו קטנה ממש מ־1.**

רמז: זה עוזר קודם להוכיח שהסיכוי שיהיה i כל שהוא שעבורו $X_i = 2$ הוא ריבוע הסיכוי שיהיה i כל שהוא שעבורו $X_i = 1$. לשם כך חישבו, עבור i נתון, מהי ההסתברות המותנה לקיום $j > i$ שעבורו $X_j = 2$ בהינתן המאורע $X_i = 1$.

הערה: מרחב ההסתברות המוגדר כאן (המרחב מעל כל הסדרות האפשריות של ערכי $X_0, X_1, X_2, \ldots$) הוא בעצם רציף. אפשר אבל לפתור את השאלה בלי להתעמק בפורמליזם של מרחבי הסתברות רציפים (בפרט כל המאורעות המעורבים יהיו "לגיטימיים").

סיכויים לרמזי

הראו שלכל מספר טבעי $k > 1$ יש קבוע $\alpha_k > 0$ עם התכונה הבאה: לכל פונקציה $f : [0,1] \times [0,1] \to \{0,1\}$ שהיא סימטרית (מתקיים $f(x,y) = f(y,x)$ לכל $x, y \in [0,1]$), ומדידה (המונח "מדידה" אומר כאן שאם X ו־Y הם משתנים מקריים ב"ת שנבחרים יוניפורמית מתוך הקטע $[0,1]$, אז ביטויים כמו "$f(X,Y) = 1$" הם מאורעות

שאפשר לנתח עבורם הסתברויות), אם $X_1,\dots,X_k$ הם מ"מ ב"ת לחלוטין הנבחרים יוניפורמית מ־$[0,1]$ (שימו לב שמרחב ההתפלגות שלנו רציף), אז לפחות לאחד משני המאורעות "לכל $1\le i<j\le k$ מתקיים $f(X_i,X_j)=0$" ו־"לכל $1\le i<j\le k$ מתקיים $f(X_i,X_j)=1$" יש הסתברות לפחות α_k.

פתרונות לתרגילים על בניה וניתוח של מרחבי הסתברות

התפלגויות מותנות – הכיוון השני

לכל $1\le i\le n$ נגדיר $m_i=\min\{p_i,q_i\}$. מכיוון שמתקיים $|p_i-q_i|=p_i+q_i-2m_i$, מתקיים גם

$$\sum_{i=1}^n m_i=\frac{1}{2}(\sum_{i=1}^n p_i+\sum_{i=1}^n q_i-\sum_{i=1}^n |p_i-q_i|)=1-\epsilon$$

עתה צריך לשים לב ששלושת הסדרות המוגדרות באופן הבא הן ווקטורי התפלגות (ז"א שהערכים כולם אי שליליים וסכומם הוא 1).

$$\begin{aligned} m_i' &= m_i/(1-\epsilon)\\ p_i' &= (p_i-m_i)/\epsilon\\ q_i' &= (q_i-m_i)/\epsilon \end{aligned}$$

מרחב ההסתברות שלנו יוגדר עתה לפי ערכי המ"מ X ו־Y המוגרלים באופן הבא: בהסתברות $1-\epsilon$ אנו נבחר גם ל־X וגם ל־Y ערך המוגרל מתוך $\{1,\dots,n\}$ לפי ווקטור ההתפלגות $m_1',\dots,m_n'$. בהסתברות הנותרת ϵ אנו נבחר באופן בלתי תלוי ל־X ערך המוגרל לפי $p_1',\dots,p_n'$, ול־Y ערך המוגרל לפי $q_1',\dots,q_n'$.

כדאי עתה לשים לב שלא קיים i עבורו גם p_i' וגם q_i' אינם אפס (מכיוון ש־m_i שווה לאחד מ־p_i,q_i), ולכן ההסתברות למאורע $X=i\wedge Y=j$ שווה בדיוק ל־m_i אם $i=j$, ושווה ל־$(p_i-m_i)(q_j-m_j)/\epsilon$ אם $i\ne j$. מכאן נובע שההסתברות של המאורע $X=Y$ היא בדיוק $1-\epsilon$.

נותר עוד להוכיח שלמשתנים X ו־Y יש את ההתפלגויות (הבלתי־מותנות) הרצויות. נראה זאת לדוגמה עבור X (עבור Y ההוכחה זהה):

$$\begin{aligned} \Pr[X=i] &= (1-\epsilon)\,m_i'+\epsilon p_i'\\ &= m_i+(p_i-m_i)\\ &= p_i \end{aligned}$$

התאמה זוגית

חישוב ישיר היה עובד כאן (עבור מרחבים בדידים). בפתרון זה נחסוך את החישובים ונשתמש בשאלה "התפלגויות מותנות" מהפרק על מרחק בין התפלגויות והשאלה "התפלגויות מותנות – הכיוון השני" מהפרק הזה.

לפי השאלה "התפלגויות מותנות – הכיוון השני", אנו יכולים למצוא התפלגות ξ מעל $S\times S$ כך שהצמצום שלה לקורדינטה הראשונה מתפלג כמו μ_1, הצמצום שלה לקורדינטה השניה מתפלג כמו μ_2, וההסתברות למאורע שהוגרל (i_1,i_2) שעבורו $i_1\ne i_2$ היא $d(\mu_1,\mu_2)$. השאלה הנ"ל אומנם מנוסחת במושגים של בניית שני מ"מ X,Y, אבל זה שקול להגדרה $\xi(i_1,i_2)=\Pr[X=i_1\wedge Y=i_2]$.

באופן דומה נבנה את η מעל $T\times T$ כך שהצמצום שלה לקורדינטה הראשונה מתפלג כמו ν_1, הצמצום שלה לקורדינטה השניה מתפלג כמו ν_2, וההסתברות למאורע שהוגרל (j_1,j_2) שעבורו $j_1\neq j_2$ היא $d(\nu_1,\nu_2)$. ננתח עתה את $\xi\times\eta$, ההתפלגות של הגרלת זוג ערכים לפי ξ והגרלה (ב"ת בהגרלה לפי ξ) של זוג ערכים לפי η. זוהי התפלגות מעל $(S\times S)\times(T\times T)$, אבל ע"י סידור מחדש של הקורדינטות אפשר להסתכל עליה כהתפלגות מעל $(S\times T)\times(S\times T)$.

עתה ננתח את ההתפלגות על ה"קורדינטה" הראשונה של זוגות ערכים $(i_1,j_1)\in S\times T$. אנחנו יודעים ש־i_1 מתפלג כמו μ_1 (זו הקורדינטה הראשונה של בחירת זוג לפי ξ), ו־j_1 מתפלג כמו ν_1 (זו הקורדינטה הראשונה של בחירת זוג לפי η). מכיוון שהגרלנו באופן ב"ת את הערך לפי ξ ואת הערך לפי η, הזוג כולו מתפלג לפי $\mu_1\times\nu_1$.

באופן דומה, ה"קורדינטה" השניה היא זוג ערכים $(i_2,j_2)\in S\times T$ שמתפלג לפי $\mu_2\times\nu_2$. לבסוף, נחסום את ההסתברות שההגרלה של $((i_1,j_1),(i_2,j_2))\in(S\times T)\times(S\times T)$ תקיים $(i_1,j_1)\neq(i_2,j_2)$. נזכיר שבהגדרה של $\xi\times\eta$ בחרנו את (i_1,i_2) לפי ξ, ולכן $\Pr[i_1\neq i_2]=d(\mu_1,\mu_2)$, ובחרנו את (j_1,j_2) לפי η, ולכן $\Pr[j_1\neq j_2]=d(\nu_1,\nu_2)$. כמו כן, קבענו את הבחירה של (i_1,i_2) באופן ב"ת בבחירה של (j_1,j_2), ולכן מתקיים:

$$\begin{aligned}\Pr_{\xi\times\eta}[i_1\neq i_2\vee j_1\neq j_2] &= 1-\Pr_\xi[i_1=i_2]\Pr_\eta[j_1=j_2]=1-(1-d(\mu_1,\mu_2))(1-d(\nu_1,\nu_2))\\ &= d(\mu_1,\mu_2)+d(\nu_1,\nu_2)-d(\mu_1,\mu_2)\cdot d(\nu_1,\nu_2)\end{aligned}$$

לפי השאלה "התפלגויות מותנות" (לא הכיוון השני), בניה זו נותנת את החסם הנדרש על $d(\mu_1\times\nu_1,\mu_2\times\nu_2)$.

סימולציה של הסתברות

נבצע סימולציה של בחירה יוניפורמית של $0\le\beta\le 1$, ונעצור את הסימולציה ברגע שהשאלה האם $\beta<\alpha$ או $\alpha\ge\beta$ היא בעלת תשובה וודאית. באופן פורמלי: נתחיל עם $\beta_0=0$ ו־$k=1$. בכל שלב (תוך שימוש ב"הטלת מטבע" חדשה בודדת), בהסתברות $\frac{1}{2}$ נקבע $\beta_k=\beta_{k-1}$ ובהסתברות $\frac{1}{2}$ נקבע $\beta_k=\beta_{k-1}+2^{-k}$. אם $\beta_k\ge\alpha$ אז נעצור מיידית ונחזיר "0". אם $\beta_k+2^{-k}<\alpha$ אז נעצור מיידית ונחזיר 1. בכל מקרה אחר נגדיל את k ב־1 ונעבור לשלב הבא.

על מנת להראות שתוחלת מספר השלבים (ולכן גם מספר הטלות המטבע) חסומה ע"י מספר קבוע, נראה שבכל שלב יש סיכוי של $\frac{1}{2}$ שהאלגוריתם יעצור. אם האלגוריתם לא עצר בשלב ה־$k-1$ (או קודם) אז לפי תנאי העצירה שנבדק שם מתקיים $\beta_{k-1}<\alpha\le\beta_{k-1}+2^{1-k}$. אם $\alpha\le\beta_{k-1}+2^{-k}$, אז נעצור בשלב הנוכחי אם ורק אם קבענו $\beta_k=\beta_{k-1}+2^{-k}$, ואם $\alpha>\beta_{k-1}+2^{-k}$ אז נעצור בשלב הנוכחי אם ורק אם קבענו $\beta_k=\beta_{k-1}$. בשני המקרים נעצור בהסתברות $\frac{1}{2}$.

עתה נחשב את הסיכוי הכולל שהאלגוריתם יעצור בסופו של דבר עם התשובה "1". אנחנו יודעים מהפסקה הקודמת שהסיכוי שהאלגוריתם לא עצר עד סוף השלב ה־k הוא בדיוק 2^{-k}, וזה כולל את המקרה "$k=0$" (האלגוריתם בסיכוי 1 יבצע את השלב הראשון). נראה באינדוקציה שאם $l2^{-k}\le\alpha<(l+1)2^{-k}$, אז בסיכוי $l2^{-k}$ בדיוק האלגוריתם יעצור עד סוף השלב ה־k עם תשובה 1. זה אומר שהסיכוי שהאלגוריתם יעצור בשלב כל שהוא עם תשובה "1" הוא α בדיוק.

נניח אם כן ש־$l'2^{1-k}\le\alpha<(l'+1)2^{1-k}$, ובאינדוקציה שהסיכוי שהאלגוריתם עצר עם "1" עד השלב ה־$k-1$ הוא $l'2^{1-k}$ בדיוק (שימו לב שההנחה נכונה עבור הבסיס $k=1$ עם $l'=0$). אנו גם יודעים שהמאורע (הזר) שהאלגוריתם הגיע לראשית השלב ה־k הוא 2^{1-k}. במקרה הנ"ל אנו יודעים גם ש־$\beta_{k-1}=l'2^{1-k}$ (ברור ש־β_{k-1} הוא כפולה שלמה של 2^{1-k}, והמכפיל הוא לפי התנאי שהאלגוריתם בדק בשלב ה־$k-1$). נותר רק לבדוק את שני המקרים: אם $\alpha\le\beta_{k-1}+2^{-k}$ אז $l=2l'$ ואכן ההתפלגות (המותנה על הגעה לשלב ה־k) שהאלגוריתם יענה "1" בדיוק בשלב זה היא אפס. אם $\alpha>\beta_{k-1}+2^{-k}$ אז $l=2l'+1$ וההתפלגות (המותנה) שהאלגוריתם יענה

"1" עתה היא $\frac{1}{2}$, אשר בחישוב הכולל תיתן תנו את הסכום המבוקש $l2^{-k}$ עבור ההסתברות לתשובה "1" בשלב כל שהוא עד סוף השלב ה־k.

התבלבלות

באופן פורמלי (פורמליזם־יתר יכול לעזור להבנה כאן), ההתפלגות $\hat{\mu} : S \times S \to [0,1]$ המוגדרת בשאלה נתונה ע"י הפונקציה $\hat{\mu}(i,j) = \frac{1}{2}(\mu_1(i)\mu_1(j) + \mu_2(i)\mu_2(j))$, והמ"מ מוגדרים ע"י הפונקציות $X(i,j) = i$ ו־$Y(i,j) = j$.

עתה נניח שקיים i שעבורו $\mu_1(i) \neq \mu_2(i)$, ונראה שהמ"מ אינם ב"ת זה בזה ע"י כך שנראה שמתקיים אי־השוויון $\Pr[X = i \wedge Y = i] \neq \Pr[X = i] \cdot \Pr[Y = i]$. מחישוב ישיר $\Pr[X = i \wedge Y = i] = \frac{1}{2}((\mu_1(i))^2 + (\mu_2(i))^2)$ וכן $\Pr[X = i] \cdot \Pr[Y = i] = (\frac{1}{2}(\mu_1(i) + \mu_2(i)))^2$. להשלמת ההוכחה מחשבים את ההפרש:

$$\Pr[X = i \wedge Y = i] - \Pr[X = i] \cdot \Pr[Y = i] = \frac{(\mu_1(i))^2}{4} + \frac{(\mu_2(i))^2}{4} - \frac{\mu_1(i)\mu_2(i)}{2} = \frac{(\mu_1(i) - \mu_2(i))^2}{4} \neq 0$$

עולה עולה

אנחנו נשתמש בהוכחה בטכניקה שמזכירה את השאלה "לבלוע את החכמה" מהפרק על לינאריות התוחלת (אבל כאן בלי שימוש בשיטה זו). אנחנו נגדיר מרחב הסתברות $\mu_{p,q}$ מעל $\mathcal{P}(S) \times \mathcal{P}(S)$ שבו מוגרל זוג של קבוצות (A,B), באופן הבא: לכל $s \in S$, באופן ב"ת בכל $t \in S$ אחר, נבחר עבורו אחת מהתוצאות הבאות: בהסתברות p נקבע $s \in A \cap B$, בהסתברות $q - p$ נקבע $s \in B \setminus A$ (כזכור לפי ההנחות שלנו $p < q$), ובהסתברות הנותרת $1 - q$ נקבע $s \notin A \cup B$. שימו לב שעבור הגרלה כזו של (A,B) מתקיימים הדברים הבאים:

- ההתפלגות (הלא־מותנה) של A זהה ל־μ_p.
- ההתפלגות (הלא־מותנה) של B זהה ל־μ_q.
- תמיד (בהסתברות 1) מתקיים $A \subseteq B$.
- בהסתברות $(q-p)^{|S|} > 0$ מתקיים $(A,B) = (\emptyset, S)$, כי $p < q$.

מהסעיף הראשון מתקיים $\Pr_{\mu_{p,q}}[A \in \mathcal{A}] = \Pr_{\mu_p}[\mathcal{A}]$ ומהסעיף השני מתקיים $\Pr_{\mu_{p,q}}[B \in \mathcal{A}] = \Pr_{\mu_q}[\mathcal{A}]$. מכיוון ש־$\mathcal{A}$ היא מונוטונית לא־יורדת, לפי הסעיף השלישי המאורע $A \in \mathcal{A}$ מוכל במאורע $B \in \mathcal{A}$, ולכן $\Pr_{\mu_p}[\mathcal{A}] \leq \Pr_{\mu_q}[\mathcal{A}]$.

הדרישה של השאלה אבל היא לאי־שוויון חזק, ולשם כך נשתמש בסעיף הרביעי. בשאלה נתון שמתקיים $\mathcal{A} \neq \mathcal{P}(S)$, ולכן בפרט $\emptyset \notin \mathcal{A}$ (כי אחרת $\mathcal{A}$ היה מכיל את כל הקבוצות המכילות את הקבוצה הריקה, ז"א את כל $\mathcal{P}(S)$). כמו כן נתון שמתקיים $\mathcal{A} \neq \emptyset$, ולכן בפרט $S \in \mathcal{A}$ (כי אם $\mathcal{A}$ מכיל קבוצה כל שהיא אז הוא מכיל גם את S המכילה אותה). על כן הסיכוי שמתקיים $B \in \mathcal{A}$ אך לא מתקיים $A \in \mathcal{A}$ הוא לפחות $(q-p)^{|S|} > 0$, ולכן בפרט $\Pr_{\mu_q}[\mathcal{A}] - \Pr_{\mu_p}[\mathcal{A}] > 0$ כנדרש.

ניחושים מול שקרן

לפני שנתאר את האלגוריתם עצמו, קצת אינטואיציה: כל אלגוריתם חייב לעצור בפעם הראשונה שהוא מקבל את התשובה "כן" (כי אז מצאנו את המספר הנכון), ולכן נותר לכתוב את סדרת השאלות שהאלגוריתם שואל כל עוד הוא לא קיבל תשובה כזו – זוהי סדרה אינסופית של מספרים טבעיים. אם יש i שמופיע מספר סופי של פעמים בסדרה הנ"ל, אז יש הסתברות חיובית לא למצוא את המספר הנכון אף פעם (במקרה ש־i הוא המספר הנכון), ולכן אין i כזה.

אנחנו נחלק את האלגוריתם לסדרה של שלבים, כשבכל שלב האלגוריתם יחזור גם על השאלות מהשלבים הקודמים. בשלב ה־l (נניח שמתחילים מ־0 $= l$) האלגוריתם ישאל את כל האיברים $\{1,\dots,a_l\}$. נחשב את התוחלת של $\sum_{l=0}^{L} a_l$, כאשר L הוא המ"מ של מספר השלב שבו קיבלנו תשובה "כן" (זה חוסם את מספר השאלות שבצענו).

אם k הוא התשובה הנכונה, ו־r הוא ה־l המינימלי שעבורו מתקיים $a_l \geq k$, אז $\Pr[L=l]$ שווה ל־0 אם $l<r$ ושווה ל־$2/3^{l+1-r}$ אם $r \geq l$. אם רוצים לחשב את התוחלת, אז נוח לחשב את ההסתברות $\Pr[L \geq l]$, ששווה ל־3^{r-l} עבור $l \geq r$ (אפשר לחשוב על זה כעל ההסתברות שלפחות $l-r$ פעמים השקרן שיקר בקשר ל־k), ושווה ל־1 עבור $l \leq r$. נוסחת התוחלת אז היא

$$\mathrm{E}[\sum_{l=0}^{L} a_l] = \sum_{l=0}^{\infty}(\sum_{t=0}^{l} a_t)\Pr[L=l] = \sum_{l=0}^{\infty} a_l \Pr[L \geq l] = \sum_{t=0}^{r} a_l + \sum_{t=1}^{\infty} 3^{-t} a_{r+t}$$

אנחנו נבחר $a_l = 2^l$, ז"א שבשלב ה־l האלגוריתם שואל את השקרן את כל 2^l השאלות החל מ־"האם המספר הוא 1?" ועד "האם המספר הוא 2^l?". הסיבה לבחירה כזו היא לדאוג לחסם קטן עבור האיבר $\sum_{t=0}^{r} a_l$ (ובגלל הדעיכה המהירה של $\Pr[L \geq l]$, גם שאר הסכום יהיה קטן). במקרה שלנו, אם התשובה היא k, אז מספר השלבים שהאלגוריתם בטוח יעשה הוא $r = \lceil \log k \rceil$. מכאן אפשר לקבל את החסם עבור תוחלת מספר השאלות הכולל:

$$\sum_{t=0}^{r} 2^t + \sum_{t=1}^{\infty} 3^{-t} 2^{r+t} = (2^{r+1}-1) + 2^r \sum_{t=1}^{\infty} (\frac{2}{3})^t = (2^{r+1}-1) + 2^r \cdot 2 < 2^{r+2} < 8k$$

וקטורים בהגרלה

יש (לפחות) שני פתרונות אפשריים.

פתרון בעזרת נוסחת ההתפלגות המותנה: עבור $0 \leq i \leq n$, מגדירים את E_i להיות המאורע ש־$\{v_1,\dots,v_i\}$ היא קבוצה ב"ת. באופן פורמלי E_0 הוא מאורע המתקיים בהסתברות 1, שקבוצת הווקטורים הריקה היא ב"ת. שימו לב שהמאורע E_i גורר את כל המאורעות $E_0,\dots,E_{i-1}$. זה אומר שכאשר מסתכלים עליהם כתתי־קבוצה, מתקיים $E_0 \supseteq \cdots \supseteq E_{i-1} \supseteq E_i$ מכך ניתן להוכיח באינדוקציה שמתקיים $\Pr[E_i] = \prod_{j=1}^{i} \Pr[E_j|E_{j-1}]$: בסיס האינדוקציה $i=1$ נובע מכך שמתקיים $\Pr[E_0]=1$. על מנת להוכיח את המעבר מ־$i-1$ ל־i כותבים $\Pr[E_i] = \Pr[E_i \wedge E_{i-1}] = \Pr[E_i|E_{i-1}]\Pr[E_{i-1}] = \prod_{j=1}^{i} \Pr[E_j|E_{j-1}]$.

ההסתברות $\Pr[E_i|E_{i-1}]$ היא $1-\Pr[w_i \in \mathrm{span}\{w_1,\dots,w_{i-1}\}]$, וע"י חישוב המימד של המרחב הנפרש על הווקטורים הקודמים (כאשר לפי E_{i-1} כבר ידוע שזו קבוצה ב"ת), מתקבל $\Pr[E_i|E_{i-1}] = 1-2^{i-1-n}$. מאלו מתקבל $\Pr[E_i] = \prod_{j=1}^{n}(1-2^{j-1-n}) \geq \prod_{j=1}^{\infty}(1-2^{-j})$, ולפי מה שאתם זוכרים (אולי) מאינפי', המכפלה האינסופית הימנית גדולה מ־0 (נובע מההתכנסות של $\sum_{j=1}^{\infty} 2^{-j}$; או אפשר להשתמש בזה שעבור כל $x>0$ קטן מספיק מתקיים $1-x > e^{-2x}$; יש עוד אפשרויות הוכחה, אבל ממילא יש גם את הפתרון האלטרנטיבי למטה).

פתרון ע"י איחוד מאורעות: ראשית מוכיחים חסם על ההסתברות ש־$\{v_1,\dots,v_{n-1}\}$ ב"ת. נסתכל על סידרת ערכים $\alpha_1,\dots,\alpha_{n-1} \in \{0,1\}$ שלא כולם אפס, ונשים לב שמתקיים $\Pr[\sum_{i=1}^{n-1}\alpha_i v_i = 0] = 2^{-n}$ (עבור ווקטורים $v_1,\dots,v_{n-1}$ הנבחרים מקרית). לפי איחוד מאורעות, מכיוון שיש $2^{n-1}-1$ סדרות אפשריות של $\alpha_1,\dots,\alpha_{n-1}$, בהסתברות לפחות $\frac{1}{2}$ אין סדרה $\alpha_1,\dots,\alpha_{n-1}$ שמאפסת את הקומבינציה הלינארית המתאימה של הווקטורים, ומכאן שקבוצת הווקטורים היא ב"ת. ע"מ לחסום את ההסתברות ש־$\{v_1,\dots,v_n\}$ כולה היא ב"ת, עושים כמו בפתרון הקודם שימוש (הפעם יחיד) בהסתברות המותנה על המאורע עבור $\{v_1,\dots,v_{n-1}\}$, ומקבלים חסם תחתון של $\frac{1}{4}$ על המאורע המבוקש.

קואליציה רחבה

לכל תת־קבוצה $I \subset \{1,\ldots,10m\}$ שעבורה $|I| \leq m$, נגדיר את B_I כמאורע ש־I "ממחיש את המקסימום של f עבור ת"ק של $\{X_1,\ldots,X_{10m}\}$", ז"א שמתקיים $f(\{X_i : i \in I\}) = \max_{J\subseteq\{1,\ldots,10m\}} f(\{X_i : i \in J\})$. נראה עתה שמתקיים $\Pr[B_I] \leq 2^{-9m}$ לכל I. לשם כך נראה שלכל סדרה $\langle a_i : i \in I\rangle$ של ערכים מ־S מתקיים $\Pr[B_I|\forall_{i\in I} X_i = a_i] \leq 2^{-9m}$, ומכך נובע המבוקש לפי נוסחת ההסתברות השלמה.

בהנתן הסדרה $\langle a_i : i \in I\rangle$, אם $f(\{X_i : i \in I\}) < \max_{J\subset I} f(\{X_i : i \in J\})$ אז בבירור מתקיים $\Pr[B_I|\forall_{i\in I} X_i = a_i] = 0$, ולכן מעתה נתמקד במקרה $f(\{X_i : i \in I\}) = \max_{J\subseteq I} f(\{X_i : i \in J\})$.

עתה נגדיר לכל $j \in \{1,\ldots,10m\} \setminus I$ את המאורע $B_{I,j}$, "אי אפשר להגדיל את הערך של I ע"י שימוש ב־j", שבאופן פורמלי מתקיים אם $f(\{X_i : i \in I\}) = \max_{J\subseteq I\cup\{j\}} f(\{X_i : i \in J\})$. נשים לב שעבור המאורעות האלו מתקיים $\Pr[B_I|\forall_{i\in I} X_i = a_i] \leq \Pr[\bigwedge_{j\in\{1,\ldots,10m\}\setminus I} B_{I,j}|\forall_{i\in I} X_i = a_i]$ (כי בפרט B_I גורר את $B_{I,j}$ לכל j). מתקיים גם $\Pr[B_{I,j}|\forall_{i\in I} X_i = a_i] \leq \frac{1}{2}$ לפי הנתון מהשאלה (עם $A = \{a_i : i \in I\}$). תחת ההתניה על הערכים של X_i עבור $i \in I$ כל המאורעות האלו הם ב"ת (כי $B_{I,j}$ תחת ההתניה הנ"ל תלוי רק ב־X_j, ונתון שהערכים של $X_1,\ldots,X_{10m}$ נבחרים באופן ב"ת לחלוטין). על כן מתקיים

$$\Pr[\bigwedge_{j\in\{1,\ldots,10m\}\setminus I} B_{I,j}|\forall_{i\in I} X_i = a_i] = \prod_{j\in\{1,\ldots,10m\}\setminus I} \Pr[B_{I,j}|\forall_{i\in I} X_i = a_i] \leq 2^{-9m}$$

עתה שהוכחנו את $\Pr[B_I|\forall_{i\in I} X_i = a_i] \leq 2^{-9m}$, נשים לב שמספר ה־$I$ האפשריים בהגדרת המאורע הוא $\sum_{k=0}^{m}\binom{10m}{k} \leq \binom{11m}{m} \leq (\frac{11em}{m})^m < 2^{8m}$, ולכן בהסתברות גדולה מ־$\frac{1}{2}$ אף B_I לא מתקיים. כאשר זה קורה, המקסימום $\max_{J\subseteq\{1,\ldots,10m\}} f(\{X_i : i \in J\})$ חייב להיות מושג ע"י קבוצה J שגודלה גדול מ־m. עבור קבוצה זו בפרט מתקיים $f(\{X_i : i \in J\}) > f(\emptyset) \geq 0$, ז"א שעבור הקבוצה J מתקבל ערך חיובי.

בחירה בלי מגוון

נבצע את ההגרלה בצורה הבאה: עבור $1 \leq i \leq c$, נסמן את $A_i = \{a : f(a) = i\}$. עתה נגריל באופן מקרי ערך J כך ש־$\Pr[J = i] = |A_i|/n$ לכל $1 \leq i \leq c$ (שימו לב שסכום ההסתברויות הוא אכן 1). בהינתן הערך של J, אם $|A_J| < k$ אז נגדיר $K = \emptyset$, ואחרת נגריל את K באופן יוניפורמי מתוך $\binom{A_J}{k}$, משפחת תתי הקבוצה של A_J מגודל k. הדברים הבאים מתקיימים עבור $a \in A$, בהתאמה לסעיפי השאלה:

- אם $|A_{f(a)}| < k$ אז $\Pr[a \in K] = 0$ (כי אם "קבוצת הצבע" של a מוגרלת, אז היא תהיה קטנה מדי, ואז בהכרח יבחר $K = \emptyset$), ואם $|A_{f(a)}| \geq k$ אז לפי שימוש בנוסחת ההסתברות המותנה מתקיים $\Pr[a \in K] = \Pr[a \in K|J = f(a)]\Pr[J = f(a)] = (k/|A_{f(a)}|)(|A_{f(a)}|/n) = k/n$.
- מכיוון שלפי ההגדרות למעלה תמיד $K \subseteq A_J$ עבור J כל שהוא, לכל $a \in K$ מתקיים $f(a) = J$, ז"א שהערכים של f שווים זה לזה עבור איברי K.
- ההסתברות לכישלון היא סכום ההסתברויות של $\Pr[J = i]$ עבור i המקיימים $|A_i| < k$, ומכיוון שיש רק c ערכים אפשריים ל־J, ההסתברות הזו חסומה ע"י $c(k-1)/n$. אפשר אם כן לקבוע $N(c,k,\epsilon) = c(k-1)/\epsilon$ לקבלת החסם הדרוש על $\Pr[K = \emptyset]$.

מרחב הסתברות בסגנון רמזי

אנחנו נתאר כאן לכל ϵ ולכל גרף דו־צדדי עם n צמתים שצבוע ב־c צבעים מרחב הסתברות μ_ϵ, כך שאם $n \geq M(c,k,\epsilon)$ (אח"כ נכתוב את החישוב של M), אז (X,Y) שנבחרים לפי μ_ϵ יקיימו את הדרוש. מספיק לספק תאור כזה מהסיבה הבאה: בכל מקרה אנחנו נדרשים לתאר מרחב μ שתלוי בגרף הדו־צדדי הנתון. אנחנו נמצא את ϵ המינימלי שעבורו $n \geq M(c,k,\epsilon)$ (אפשר להניח ש־M מונוטונית ב־ϵ), ונקבע $\mu = \mu_\epsilon$.

הבחירה שלנו תהיה לפי השלבים הבאים. הפונקציה N המופיעה בהם היא זו מהשאלה "בחירה בלי מגוון".

1. נבחר באופן יוניפורמי קבוצה $X' \subset U$ מתוך $\binom{U}{k'}$, משפחת כל תתי־הקבוצה של U מגודל k'. אנחנו נשתמש ב־$k' = N(c,k,\epsilon/2)$: הבחירה הזו תוסבר אח"כ (בינתיים כדאי לחשוב על k' כ"מספיק גדול בשביל ההמשך"). נסמן $X' = \{x'_1, \ldots, x'_{k'}\}$, לפי סדר שהגדרנו מראש על כל צמתי U. אח"כ נבחר את $M(c,k,\epsilon)$ כך שיהיה גדול מ־k', כך שבחירה כזו של X' תהיה אפשרית עבור $n \geq M$ (ועבור $n < k'$ אפשר פשוט להחזיר $(\emptyset,\emptyset)$ תמיד).
2. נצבע את V ב־$c^{k'}$ צבעים: לכל $v \in V$, נגדיר את $f'(v)$ להיות הסדרה $f(x'_1 v), \ldots, f(x'_{k'} v)$ (יש סה"כ $c^{k'}$ אפשרויות).
3. נשתמש בתוצאת השאלה "בחירה בלי מגוון" לבחירת קבוצה Y מגודל k מתוך V. כאן המקום להגדיר את $M(c,k,\epsilon) = N(c^{k'},k,\epsilon/2)$, כך שהסיכוי לכישלון בשלב זה אינו עולה על $\epsilon/2$ (נשים לב שבפרט מתקיים כאן $M(c,k,\epsilon) > k'$). אם נכשלנו בשלב זה אז נעצור את האלגוריתם ונחזיר את הזוג $(\emptyset,\emptyset)$.
4. נשים לב שעתה לכל צומת $x' \in X'$ מתקיים שערכי $f(x',y)$ זהים זה לזה עבור כל האפשרויות של $y \in Y$, בגלל הצורה שבחרנו את Y. נתייחס אם כן ל־f כאל צביעה של צמתי X', ונשתמש שוב בשאלה "בחירה בלי מגוון" על מנת לבחור קבוצה $X \subset X'$ מגודל k. לפי הגדרת k', הסיכוי לכישלון גם בשלב זה אינו עולה על $\epsilon/2$. אם נכשלנו אז נחזיר את הזוג $(\emptyset,\emptyset)$, ואחרת נחזיר את (X,Y).

עתה ננתח את התוצאה שקיבלנו. כפי שראינו מתיאור האלגוריתם, ההסתברות הכוללת לכישלון (החזרת קבוצות ריקות) חסומה ע"י ϵ, ולפי ניתוח מצב הצביעה בשלב האחרון, כל הזוגות מתוך $X \times Y$ אכן צבועים באותו צבע. נשאר לנו עבור $(u,v) \in U \times V$ לחסום את הסיכוי שבחרנו את שני הצמתים, לפי נוסחת ההסתברות המותנה:

$$\Pr[u \in X \wedge v \in Y] = \Pr[u \in X'] \cdot \Pr[v \in Y | u \in X'] \cdot \Pr[u \in X | u \in X' \wedge v \in Y]$$

המוכפל השמאלי שווה ל־k'/n (בחירה יוניפורמית פשוטה של תת־קבוצה), המוכפל האמצעי, בגלל אופן השימוש בתוצאת השאלה "בחירה בלי מגוון", חסום ע"י k/n, והמוכפל הימני (גם בגלל השימוש ב"בחירה בלי מגוון") חסום ע"י k/k'. סה"כ קיבלנו חסם כולל של k^2/n^2 כנדרש.

דו־קרב תרנגולים

הפתרון היותר אלגנטי משתמש בהסתברות מותנה (האופציה השניה היא פשוט לחשב סכום של טור חזקות). המשחק נעצר לאחר סיבוב שיש בו ניקור אחד או יותר, ומכיוון שיש הסתברות קבועה כזו לכל סיבוב, בהסתברות 1 יהיה לבסוף סיבוב כזה. לכל j נרצה לדעת את ההסתברות שנשאר תרנגול בריא בסוף סיבוב זה, כאשר מתנים אותה על המאורע שזהו הסיבוב הראשון (ולכן היחידי) שבוצע בו ניקור. ההסתברויות לאפשרויות בסיבוב ה־j אינן תלויות בסיבובים הקודמים, ז"א שכל אחת מארבעת האפשרויות לפעולות של שני התרנגולים קורית בהסתברות $1/4$. עם זאת, אנחנו מתנים עתה על כך שהיה ניקור, ז"א שיש לנו שלוש אפשרויות שוות הסתברות, שבשתיים מתוכן נשאר תרנגול בריא. על כן קיבלנו סיכוי של $2/3$ שנותר תרנגול כזה, ומכיוון שזה נכון לכל התנייה על j, זוהי גם ההסתברות הלא־מותנה של המאורע.

דו־קרב תרנגולים עייפים

הפתרון דומה לפתרון עבור השאלה "דו־קרב תרנגולים", רק שצריך ניתוח דו־שלבי. ראשית ננתח את ההסתברויות המותנות עבור הסיבוב הראשון שבו לפחות אחד התרנגולים עושה משהו (מנקר או הולך לישון). ההסתברות ששני התרנגולים הולכים לישון בו־זמנית הוא $1/9$, אבל אם מתנים אותה על ההסתברות שלפחות אחד התרנגולים

עושה משהו (מאורע בהסתברות $8/9$) אז מקבלים הסתברות מותנה של $1/8$. כמו כן, בהסתברות מותנה של $1/4$, בסיבוב הראשון שבו נעשית פעולה, אחד התרנגולים ילך לישון והתרנגול השני לא יעשה כלום. שאר המאורעות האפשריים עבור הסיבוב הראשון שבו נעשית פעולה מסתיימים בלפחות תרנגול אחד מנוקר, ולכן אלו לא יתרמו להסתברות שאנחנו מחשבים.

אם בסיבוב הראשון שבו נעשתה פעולה נשארנו עם תרנגול אחד ערני ותרנגול אחד ישן, אז ננתח את הסיבוב הבא שבו נעשית פעולה. כאן ההסתברות המותנה שהתרנגול הערני הלך לישון היא $1/2$ (והאפשרות השניה היא שהתרנגול הערני ניקר את התרנגול הישן). סה"כ ההסתברות שבסוף כל התהליך נשארנו עם שני תרנגולים בריאים ישנים היא $1/8 + 1/4 \cdot 1/2 = 1/4$.

רגע לפני

לשם המחשה התשובה כאן היא יותר פורמלית מהמקובל. אם מתקיים $|S_0| = 1$ אז בהסתברות 1 קיים k כנדרש, ולכן נניח מעתה שמתקיים $|S_0| > 1$. ההסתברות שלא קיים k שעבורו $|S_k| \leq 1$ היא 0 (ליתר דיוק, לכל k התוחלת של $|S_k|$ היא $|S_0|p^k$ וזו שואפת ל־0 עם k, ולכן בהסתברות 1 יהיה k שעבורו $S_k = \emptyset$). זה אומר שההסתברות שלא קיים k שעבורו $|S_k| = 1$ זהה להסתברות שקיימים $k \in \mathbb{N}$ ו־$j \in \mathbb{N} \setminus \{1\}$ שעבורם $|S_{k-1}| = j$ וגם $S_k = \emptyset$, כי האופציה הנותרת היא מאורע בהסתברות 0.

כל אלו הם מאורעות זרים זה לזה, ולכן ההסתברות שלא קיים אף ערך של k שעבורו $|S_k| = 1$ זהה לסכום $\sum_{k=1}^{\infty} \sum_{j=2}^{\infty} \Pr[(S_k = \emptyset) \wedge (|S_{k-1}| = j)] = \sum_{k=1}^{\infty} \sum_{j=2}^{\infty} \Pr[(S_k = \emptyset) \wedge (|S_k| \leq 1) \wedge (|S_{k-1}| = j)]$, וזה שווה ל־$\sum_{k=1}^{\infty} \sum_{j=2}^{\infty} \Pr\left[(S_k = \emptyset) \,\middle|\, (|S_k| \leq 1) \wedge (|S_{k-1}| = j)\right] \Pr[(|S_k| \leq 1) \wedge (|S_{k-1}| = j)]$. עתה נשים לב שהמאורעות $(|S_k| \leq 1) \wedge (|S_{k-1}| = j)$ גם זרים זה לזה ולכן סכום ההסתברויות שלהם מקיים $\sum_{k=1}^{\infty} \sum_{j=2}^{\infty} \Pr[(|S_k| \leq 1) \wedge (|S_{k-1}| = j)] \leq 1$. כל שנותר הוא לחסום לכל k ו־j את ההסתברות המותנה $\Pr\left[(S_k = \emptyset) \,\middle|\, (|S_k| \leq 1) \wedge (|S_{k-1}| = j)\right]$.

לשם החסם, נחשב באופן ישיר את ההסתברות $\Pr[(S_k = \emptyset) \,|\, (|S_{k-1}| = j)] = (1-p)^j$ ואת ההסתברות $\Pr\left[(|S_k| = 1) \,\middle|\, (|S_{k-1}| = j)\right] = jp(1-p)^{j-1}$. משני אלו ניתן עתה להסיק את חסם ההסתברות המותנה $\Pr\left[(S_k = \emptyset) \,\middle|\, (|S_k| \leq 1) \wedge (|S_{k-1}| = j)\right] = (1-p)^j/((1-p)^j + jp(1-p)^{j-1}) = \frac{1-p}{1+(j-1)p} \leq \frac{1-p}{1+p}$.

לכן הסיכוי לאי קיום k שעבורו $|S_k| = 1$ חסום ע"י $\frac{1-p}{1+p} \sum_{k=1}^{\infty} \sum_{j=2}^{\infty} \Pr[(|S_k| \leq 1) \wedge (|S_{k-1}| = j)] \leq \frac{1-p}{1+p}$. בהתאמה, הסיכוי שכן קיים k כזה הוא לפחות $1 - \frac{1-p}{1+p} = \frac{2p}{1+p}$.

בלי אורך רוח

לכל $1 \leq m \leq n$ ולכל $1 \leq i \leq n+1-m$ נסמן ב־$A_{i,m}$ את המאורע $x_i = \cdots = x_{i+m-1} = 0$. נסמן ב־$Y$ את המ"מ שמקבל 0 אם $X \leq \lceil \log n \rceil$ ואחרת מקבל את הערך $X - \lceil \log n \rceil$. מספיק אם כן להוכיח שמתקיים $\mathrm{E}[Y] = O(1)$ בגלל שתמיד מתקיים $X \leq Y + \lceil \log n \rceil \leq \log n + 1 + Y$.

נפתח: $\mathrm{E}[Y] = \sum_{i=1}^{\infty} i \cdot \Pr[Y = i] = \sum_{k=1}^{\infty} \Pr[Y \geq k] = \sum_{k=1}^{n-\lceil \log n \rceil} \Pr[\bigvee_{i=1}^{n+1-k-\lceil \log n \rceil} A_{i,k+\lceil \log n \rceil}]$. בשביל ההמשך נחשב באופן ישיר $\Pr[A_{i,k+\lceil \log n \rceil}] = 2^{-k-\lceil \log n \rceil} \leq 2^{-k}/n$. לכל $1 \leq k \leq n - \lceil \log n \rceil$ מתקיים מכך לפי איחוד מאורעות $\Pr[Y \geq k] \leq \sum_{i=1}^{n+1-k-\lceil \log n \rceil} \Pr[A_{i,k+\lceil \log n \rceil}] \leq 2^{-k}$, ולכן מתקיים $\mathrm{E}[Y] \leq \sum_{k=1}^{\infty} 2^{-k} = O(1)$ כנדרש.

עם אורך רוח

נסמן ב־Z את המ"מ שמקבל 0 אם $X \geq \lfloor \log n \rfloor$ ואחרת מקבל את הערך $\lfloor \log n \rfloor - X$. בדומה לטיעון בפתרון השאלה הקודמת, כאן צריך להוכיח שמתקיים $\mathrm{E}[Z] = O(1)$. עתה נתאר שיטה אלטרנטיבית להגריל את x: נגריל מ"מ $W_1, W_2, \ldots, W_n$ באופן ב"ת לחלוטין, כאשר כל W_j מתפלג באופן גיאומטרי עם פרמטר $\frac{1}{2}$, ז"א שמתקיים

$\Pr[W_j = k] = 2^{-k}$. ראשית נבנה מחרוזת x' ע"י שרשור המחרוזות $y_1, \ldots, y_n$ כאשר y_j מורכב מ־$W_j - 1$ אפסים ולאחריו אחד בודד, ובסוף נקבע את x להיות מורכבת מ־n הביטים הראשונים של x'. הדבר לשים לב הוא ש־x שהוגרלה כאן גם מתפלגת יוניפורמית מעל $\{0,1\}^n$ (ראו את ההוכחה בסוף הפתרון), ולכן אפשר לנתח את טענת השאלה תוך שימוש במ"מ $W_1, \ldots, W_n$.

עתה נקבע את המאורע B שיש ב־x פחות מ־$n/4$ אחדות. כזכור (ראו את נוסח השאלה) מתקיים $\Pr[B] \le e^{-\Omega(n)}$. בנוסף לכך נקבע לכל $0 \le k \le \lfloor \log n \rfloor - 1$ את המאורע B_k שקובע ש־$W_1, W_2, \ldots, W_{\lceil n/4 \rceil}$ כולם קטנים מ־$\lfloor \log n \rfloor - k$. הנקודה החשובה לשים לב כאן היא שהמאורע $Z > k$ מוכל באיחוד $B_k \vee B$. כמו כן, מאי התלות של $W_1, W_2, \ldots, W_{\lceil n/4 \rceil}$, מתקיים $\Pr[B_k] = (1 - 2^{k - \lfloor \log n \rfloor})^{\lceil n/4 \rceil} < \exp(-\lceil n/4 \rceil 2^{k - \lfloor \log n \rfloor}) \le \exp(-2^{k-2})$.

מכל אלו מתקיים

$$
\begin{aligned}
\mathrm{E}[Z] &= \sum_{j=1}^{\infty} j \cdot \Pr[Z = j] = \sum_{k=0}^{\lfloor \log n \rfloor - 1} \Pr[Z > k] \le \sum_{k=0}^{\lfloor \log n \rfloor - 1} (\Pr[B] + \Pr[B_k]) \\
&\le \log n \cdot e^{-\Omega(n)} + \sum_{k=0}^{\lfloor \log n \rfloor - 1} \exp(-2^{k-2}) = O(1)
\end{aligned}
$$

הוכחה פורמלית של שקילות ההגרלה האלטרנטיבית: נוכיח ששיטת ההגרלה האלטרנטיבית של $x = (x_1, \ldots, x_n)$ (דרך $W_1, \ldots, W_n$) גם מתפלגת יוניפורמית מעל $\{0,1\}^n$. ההוכחה היא באינדוקציה שלמה על n. הבסיס הוא $n = 0$ והוא טריביאלי. עבור המעבר, נניח שהטענה מתקיימת לכל $0 \le m < n$ ונוכיח אותה עבור n. מקרה אחד, זה שבו $x_1 = x_2 = \cdots = x_n = 0$ (ז"א ש־x הוא כולו אפסים), נובע ישירות מכך שההסתברות עבור $W_1 \ge n$ היא אכן 2^{-n}.

לכל x אחר, נסמן את $1 \le i \le n$ שעבורו $x_i = 1$ אבל $x_1 = x_2 = \cdots = x_{i-1} = 0$ (ז"א שה־1 הכי "שמאלי" הוא במקום ה־i). נשים לב ש־x כזה יכול לקרות רק אם $W_1 = i$, דבר הקורה בהסתברות 2^{-i} בדיוק. עתה נפעיל את הנחת האינדוקציה עבור $z = (x_{i+1}, \ldots, x_n) \in \{0,1\}^{n-i}$ עם $m = n - i$ והמ"מ $W_2, \ldots, W_{n+1-i}$. לפי הנחת האינדוקציה אותו z מתקבל בהסתברות 2^{i-n} (גם כשמתנים על $W_1 = i$ בגלל אי התלות של $W_1, \ldots, W_n$). הכפלה בהסתברות עבור $W_1 = i$ נותנת הסתברות של 2^{-n} עבור x כולו.

לתפוס את הרכבת

נתחיל מתובנה מאוד שימושית לפתרון שיוצג כאן: עבור כל $i, k \in \mathbb{N}$ ו־$a_0, \ldots, a_k$, ההסתברות המותנה $\Pr[X_{i+1} = a_1, \ldots, X_{i+k} = a_k | X_i = a_0]$ זהה להסתברות $\Pr[X_1 = a_1 - a_0, \ldots, X_k = a_k - a_0]$ (הסתברות לא מותנה). הדבר נובע מחישוב ישיר: אם קיים l שעבורו $|a_l - a_{l-1}| \neq 1$ אז ההסתברות היא אפס, ואחרת היא שווה ל־$(\frac{1}{3})^r (\frac{2}{3})^{k-r}$ כאשר r הוא גודל הקבוצה $\{1 \le l \le k : a_l = a_{l-1} + 1\}$.

עתה נסמן את המאורעות הבאים: A הוא המאורע שקיים i שעבורו $X_i = 1$ (שאת ההסתברות שלו אנחנו מעונינים לחשב), B הוא המאורע שקיים j שעבורו $X_j = 2$, ולכל i נגדיר את המאורע B_i שקיים $j > i$ שעבורו $X_j = 2$ (בפרט המאורעות B_0 ו־B_1 זהים ל־B עצמו). כמו כן נסמן ב־T את המשתנה המקרי שערכו הוא ה־i הקטן ביותר שקיבלנו עבורו $X_i = 1$. חשוב לציין ש־T יכול לקבל גם את הערך "∞", הדבר קורה בדיוק כאשר המאורע A לא מתקיים.

הדבר הראשון לשים לב הוא שמתקיים $\Pr[A | X_1 = -1] = \Pr[B]$. זאת מכיוון שלפי התובנה למעלה, ההתפלגות של $X_2, X_3, \ldots$ בהינתן המאורע $X_1 = -1$ זהה להתפלגות הלא־מותנה של סדרת המ"מ $X_1 - 1, X_2 - 1, \ldots$. אם רוצים להיות יותר פורמלים, התובנה למעלה נכתבה עבור סדרות סופיות, ולכן הקביעה היא שההסתברות שלפחות אחד מ־$X_2, \ldots, X_k$ שווה ל־1 בהינתן המאורע $X_1 = -1$ זהה להסתברות הלא־מותנה שלפחות אחד מ־$X_1, \ldots, X_{k-1}$ שווה ל־2, ומכאן אפשר לקחת את הגבול של סדרת ההסתברויות.

מניתוח ההסתברויות לערכי X_1 נקבל $\Pr[A] = \frac{1}{3}\Pr[A|X_1 = 1] + \frac{2}{3}\Pr[A|X_1 = -1] = \frac{1}{3} + \frac{2}{3}\Pr[B]$. עתה ננסה למצוא בטוי אלטרנטיבי עבור $\Pr[B]$.

לשם כך נפצל את המאורע B לאיחוד זר של מאורעות, לפי החיתוכים עם המאורעות על הערכים האפשריים של T. ההסתברות היא חיבורית מעל מאורעות זרים, ולכן נקבל $\Pr[B] = \sum_{t\in\mathbb{N}} \Pr[B \wedge (T = t)]$ (נשים לב שמתקיים $\Pr[B \wedge (T = \infty)] = 0$ ולכן הוא לא נכנס לסכום). עבור מחובר בודד, הדבר הבא הוא לשים לב שהמאורע $B \wedge (T = t)$ זהה למאורע $B_t \wedge (T = t)$, בגלל שעבור $i \le T$ לא יכול להתקיים $X_i = 2$ (בשביל "להגיע לערך 2" צריך קודם "לעבור דרך 1", ולפי ההגדרה של T זה לא יכול לקרות לפניו). על כן מתקיים $\Pr[B \wedge (T = t)] = \Pr[B_t \wedge (T = t)] = \Pr[B_t|T = t]\Pr[T = t]$ לכל $t \in \mathbb{N}$.

עתה נשתמש שוב בתובנה השימושית מתחילת הפתרון. הפעם נקבל ממנה שההסתברות המותנה $\Pr[B_t|T = t]$ זהה להסתברות הלא־מותנה $\Pr[A]$. על כן:

$$\Pr[B] = \sum_{t\in\mathbb{N}} \Pr[B_t|T = t]\Pr[T = t] = \Pr[A] \sum_{t\in\mathbb{N}} \Pr[T = t] = (\Pr[A])^2$$

לסיום, נציב את זה במשוואה הקודמת שקיבלנו, ונקבל $\Pr[A] = \frac{1}{3} + \frac{2}{3}(\Pr[A])^2$. פתרונות המשוואה הריבועית הזו הם $\frac{1}{2}$ ו־1. היה נתון לכם שמתקיים $\Pr[A] < 1$, מה שמשאיר לנו את האפשרות $\Pr[A] = \frac{1}{2}$.

סיכויים לרמזי

נסמן ב־$R(k)$ את מספר הצמתים המינימלי שלגביו משפט רמזי מתקיים (כך שבכל גרף בעל $R(k)$ צמתים יש או קליק בעל k צמתים או קבוצה חסרת קשתות בעלת k צמתים), ונסתכל על התהליך הבא לבחירת $X_1, \ldots, X_k$: ראשית נגריל $R(k)$ מ"מ יוניפורמים מתוך $[0, 1]$ וב"ת זה בזה, $Y_1, \ldots, Y_{R(k)}$. אחר זאת נגריל סדרה של אינדכסים $1 \le i_1 < i_2 < \cdots < i_k \le R(k)$ באופן יוניפורמי מבין $\binom{R(k)}{k}$ האפשרויות (בעצם מגרילים באופן יוניפורמי ת"ק של $\{1, \ldots, R(k)\}$ בגודל k ואז מסדרים אותה לקבלת האינדכסים). לבסוף לכל $1 \le j \le k$ נקבע $X_j = Y_{i_j}$.

ראשית, נשים לב שגם בתהליך הזה, $X_1, \ldots, X_k$ מתפלגים בדיוק כמו k מ"מ ב"ת שמוגרלים יוניפורמית מ־$[0, 1]$. הסיבה לכך היא שלכל סדרה ספציפית אפשרית $1 \le i_1 < i_2 < \cdots < i_k \le R(k)$, סדרת המ"מ $Y_{i_1}, \ldots, Y_{i_k}$ מתפלגת כמו סדרה של k מ"מ ב"ת יוניפורמים, ולכן הדבר נכון גם אם השתמשנו בהגרלה כל שהיא על מנת לבחור את סדרת האינדכסים (חשוב אבל ששיטת הבחירה אינה תלויה בערכים $Y_1, \ldots, Y_{R(k)}$ עצמם, ז"א שהיינו יכולים להגריל את ה־Y_i לאחר קביעת סדרת האינדכסים).

עתה "נעקם" טיפה את הסימונים ונסמן למשל ב־$f(X_1, \ldots, X_k)$ את הגרף מעל $\{1, \ldots, k\}$ שעבורו i, j היא קשת אם ורק אם $f(X_i, X_j) = 1$. עבור גרף G עם קבוצת הצמתים $\{1, \ldots, R(k)\}$, ניתן לנתח את המאורע "$f(Y_1, \ldots, Y_{R(k)}) = G$" וגם לנתח הסתברויות שמותנות עליו, בגלל הנתון ש־f היא מדידה (המאורע הנ"ל הוא חיתוך של מאורעות מהצורה "$f(X_i, X_j) = 1$" ומשלימים שלהם).

לכל גרף G קיימת לפחות קבוצה אחת של צמתים שהיא או קליק או חסרת קשתות, לפי משפט רמזי. ההסתברות שבדיוק האינדכסים של קבוצה זו נבחרו עבור $i_1, \ldots, i_k$ היא $1/\binom{R(k)}{k}$. על כן, בהסתברות לפחות $1/\binom{R(k)}{k}$ יתקיים שהגרף $f(X_1, \ldots, X_k) = f(Y_{i_1}, \ldots, Y_{i_k})$ הוא או חסר קשתות או קליק, לפי נוסחת ההסתברות השלמה (כי החסם מתקיים עבור ההתניה על המאורע "$f(Y_1, \ldots, Y_{R(k)}) = G$" לכל גרף G אפשרי בעל $R(k)$ צמתים).

על מנת לסיים את פתרון השאלה נציב $\alpha_k = 1/2\binom{R(k)}{k}$, כי הדבר אומר שלפחות אחת משתי האפשרויות עבור $f(X_1, \ldots, X_k)$ חייבת להתקיים לפחות בהסתברות הזו (אחרת נקבל סתירה לפי אי השוויון על איחוד מאורעות).

השיטה הבסיסית

חיפוש סופה של רשימה מקושרת

נתונה רשימה מקושרת (linked list) כך שלאיברי הרשימה יש אינדקסים המסודרים בסדר עולה, וכך שיש לאלגוריתם גם את האפשרות לבחירה מקרית של איבר מהרשימה. נסתכל על האלגוריתם הבא למציאת האיבר האחרון:

מתחילים מהאיבר הראשון. בכל שלב בוחרים איבר אקראי בהסתברות אחידה, ומשווים את האינדקס שלו לאינדקס של האיבר העוקב לאיבר שנבחר בשלב הקודם. בוחרים את האיבר בעל האינדקס הגבוה יותר מביניהם עבור השלב הבא. עוצרים כאשר האיבר שנבחר הוא האחרון (אין לו איבר עוקב).

א. הראו שבהסתברות לפחות $\frac{1}{2}$ האלגוריתם יעצור בזמן $O(\sqrt{n})$; הסיקו מכך שתוחלת זמן הריצה היא $O(\sqrt{n})$, כאשר n מסמן את אורך הרשימה.

ב. הראו שתוחלת זמן הריצה של האלגוריתם היא $\Omega(\sqrt{n})$.

אוטומורפיזמים בגרף מקרי

הראו שלגרף המקרי $G(n, \frac{1}{2})$ אין אוטומורפיזם לא טריביאלי בהסתברות $1 - o(1)$. ז"א, שבהסתברות $1 - o(1)$, לכל פרמוטציה של הצמתים של G שאינה פונקצית הזהות תהיה קשת שתעבור לזוג צמתים שאינו קשת (או זוג שאינו קשת שיעבור לקשת).

הכל קשור

הראו שהדבר הבא מתקיים לכל $\alpha > 1$ קבוע: אם נגריל גרף מעל n צמתים לפי המרחב $G(n, \alpha \ln(n)/n)$ (שימו לב ללוגריתם בבסיס טבעי שם), אז הגרף יהיה קשיר בהסתברות $1 - o(1)$ (ז"א שההסתברות לאי־קשירות שואפת לאפס כאשר n שואף לאינסוף).

רמז: שימוש בהגדרה הישירה של קשירות לפי המצאות מסלול בין כל שתי נקודות היא **לא** השיטה המומלצת לפתור את השאלה.

פגיעות במרחב

נבחר תת־קבוצה A של המרחב הלינארי $(\mathbb{Z}_2)^n$, המרחב ממימד n מעל השדה הסופי בעל שני האיברים, בצורה הבאה: כל וקטור $v \in (\mathbb{Z}_2)^n$ ייבחר להיות איבר ב־A בהסתברות $2n/2^n$, באופן ב"ת בבחירות עבור כל הווקטורים האחרים. הראו שבהסתברות $1 - o(1)$ הקבוצה A פורשת את כל המרחב.

הולכים במעגלים

נתונה פרמוטציה $\sigma : \{1, \ldots n\} \to \{1, \ldots, n\}$. כידוע כל פרמוטציה ניתנת לפירוק באופן יחיד למעגלים זרים. לכם נתון שלפרמוטציה σ יש לפחות k מעגלים זרים בפירוק. כתבו (והוכיחו) אלגוריתם שבהסתברות לפחות $\frac{1}{2}$ מוצא את האינדכסים של מעגל שלם של σ (גם אינדכס בודד v שעבורו $\sigma(v) = v$ הוא "מעגל" לגיטימי). על האלגוריתם לעשות זאת תוך כדי שהוא קורא לא יותר מ־$O(1 + (n/k)(\log(n/k))^2)$ ערכים של σ במהלך הריצה (זמן הריצה שלו גם יהיה כזה, אם מניחים שקריאת ערך ועדכון מונה הן פעולות שלוקחות $O(1)$).

הדרכה: כדאי בשלב ראשון לסמן ב־m_j את מספר המעגלים בפירוק שגודלם בין 2^j ל־$2^{j+1} - 1$.

קשה לכסות

נסתכל על ה"קוביה" $\{1,\ldots,n\}^k$. נניח ש־$A_1,\ldots,A_m$ הן תתי־קוביה "קטנות" של $\{1,\ldots,n\}^k$. ז"א שלכל i מתקיים $A_i = A_i^{(1)} \times \cdots \times A_i^{(k)}$, כאשר כל $A_i^{(j)} \subset \{1,\ldots,n\}$ היא קבוצה חלקית ממש (לא שווה ל־$\{1,\ldots,n\}$). כמו כן נניח ש־$A_1,\ldots,A_m$ כולן זרות (זה **לא** אומר שעבור j קבוע הקבוצות $A_1^{(j)},\ldots,A_m^{(j)}$ הן זרות!). הראו שאם $m < 2^k$ אז $\bigcup_{i=1}^m A_i \neq \{1,\ldots n\}^k$.

הדרכה: מגרילים באופן יוניפורמי וב"ת תתי קבוצות $B^{(j)} \subseteq \{1,\ldots,n\}$ תחת התנאי ש־$|B^{(j)}|$ הוא אי־זוגי. מנתחים את זוגיות החיתוך של $B = \prod_{j=1}^k B^{(j)}$ עם $\bigcup_{i=1}^m A_i$.

הערה: יש דוגמה נגדית עבור A_i לא זרות, כבר במקרה $k=2$ ו־$n=3$. אתם מוזמנים לנסות למצוא אותה.

גרף מקרי ויחיד

נגדיר את משפחת ההתפלגויות על גרפים אינסופיים $G(\mathbb{N},p)$ באופן הבא - קבוצת הצמתים של הגרף היא $\mathbb{N}$, ולכל $i<j\in\mathbb{N}$, קשת (לא מכוונת) תחבר ביניהם בהסתברות p באופן ב"ת בזוגות האחרים. הראו כי שני גרפים G ו־H הנבחרים באופן ב"ת לפי ההתפלגות $G(\mathbb{N},\frac{1}{2})$ הם איזומורפיים בהסתברות 1. איזומורפיזם כאן הוא פונקציה חח"ע ועל בין קבוצות הצמתים (האינסופיות) שמקיימת את התנאי הרגיל ש־u,v היא קשת אם ורק אם $f(u),f(v)$ היא קשת.

פתרונות לתרגילים על השיטה הבסיסית

חיפוש סופה של רשימה מקושרת

א. נראה שבסיכוי לפחות $\frac{1}{2}$ האלגוריתם יעצור לאחר $2\sqrt{n}$ שלבים. תהי A קבוצת $\sqrt{n}$ האיברים האחרונים ברשימה המקושרת. הסיכוי שבלפחות אחד מ־$\sqrt{n}$ השלבים הראשונים האלגוריתם בחר איבר מ־A (בכל שלב האלגוריתם מבצע גם מעקב אחרי הרשימה וגם בחירה אקראית של איבר) הוא לפחות $1-\left(1-\frac{1}{\sqrt{n}}\right)^{\sqrt{n}} > \frac{1}{2}$ עבור n גדול דיו. במידה ואיבר מ־A נבחר, האלגוריתם יעצור לאחר לא יותר מ־$\sqrt{n}$ שלבים נוספים, ומכאן החסם.

באשר לתוחלת זמן הריצה, נחזור על החישוב עבור הסיכוי שהאלגוריתם יעצור לאחר יותר מ $2k\sqrt{n}$ שלבים. האלגוריתם יעצור לאחר $2k\sqrt{n}$ שלבים לכל היותר אם באחד מ־$(2k-1)\sqrt{n}$ השלבים הראשונים יבחר איבר מ־$\sqrt{n}$ האיברים האחרונים, וכך ההסתברות שנעצור לאחר $2k\sqrt{n}$ שלבים לכל היותר חסומה מלמטה על ידי $1-\left(1-\frac{1}{\sqrt{n}}\right)^{(2k-1)\sqrt{n}} > 1-e^{-(2k-1)} > 1-2^{-k}$. לכן התוחלת חסומה ע"י $\sum_{k=1}^{\infty} 2^{-k} 2k\sqrt{n} = 4\sqrt{n}$.

לחישוב הסכום הנ"ל (וסכומים דומים) כדאי לדעת את השוויון השימושי הבא המוכח ע"י חילוף משתנים בסכום: $\sum_{i=1}^{\infty} i\cdot\alpha_i = \sum_{i=1}^{\infty}\left(\sum_{j=1}^{i}\alpha_i\right) = \sum_{j=1}^{\infty}\left(\sum_{i=j}^{\infty}\alpha_i\right)$. למשל נובע מזה שמתקיים למ"מ X המקבל ערכים טבעיים בלבד השוויון $\mathrm{E}[X] = \sum_{i=1}^{\infty} i\cdot\Pr[X=i] = \sum_{j=1}^{\infty}\Pr[X\geq j]$.

ב. נראה שבסיכוי לא יותר מ־$\frac{1}{2}$ האלגוריתם יעצור לאחר $\frac{1}{2}\sqrt{n}$ שלבים לכל היותר; מכך נובע שתוחלת זמן הריצה היא לפחות $\frac{1}{4}\sqrt{n}$. נסמן את A כמקודם. הסיכוי שהאלגוריתם בחר איבר מ־A באחד מ־$\frac{1}{2}\sqrt{n}$ השלבים הראשונים הוא בוודאי לא יותר מ־$\frac{1}{2}$ (פשוט חוסמים את סכום הסיכויים למציאת איבר כזה). אם איבר כזה לא נבחר, אז האלגוריתם לא יעצור בשלבים אלו: אם נסמן ב־i את השלב האחרון שבו האיבר הנבחר אקראית היה בעל אינדקס גדול מהאיבר העוקב לזה של השלב הקודם, אז מכך שאיבר זה אינו ב־A נובע שהאלגוריתם לא היה יכול להגיע עד סוף הרשימה בשלבים הנותרים עד $\frac{1}{2}\sqrt{n}$.

אוטומורפיזמים בגרף מקרי

לכל פרמוטציה $\sigma : V \to V$ שאינה הזהות, נסמן ב־E_σ את המאורע שהיא אוטומורפיזם של הגרף הנבחר G, ואז נראה שסכום ההסתברויות על כל הפרמוטציות האפשריות הוא $o(1)$ כנדרש. אם נסמן ב־$k(\sigma)$ את מספר זוגות הצמתים של G שהפרמוטציה "מזיזה" (ז"א מספר הזוגות $\{u, v\}$ עבורם $\{u, v\} \neq \{\sigma(u), \sigma(v)\}$), אז נשים לב שמתקיים $\Pr[E_\sigma] \leq 2^{-k(\sigma)/2}$. הסיבה לכך היא שאפשר לקחת $l = k(\sigma)/2$ זוגות $\{u_1, v_1\}, \ldots, \{u_l, v_l\}$ כך שאף זוג לא מועבר ע"י σ לא לעצמו ולא לאף זוג אחר (אפשר לעשות זאת ע"י אלגוריתם חמדן), ואז המאורע ש־$\{u_i, v_i\}$ נמצא באותו סטטוס (קשת/לא־קשת) כמו $\{\sigma(u_i), \sigma(v_i)\}$ הוא בלתי תלוי במאורעות המקבילים לכל $j \neq i$. השלב הבא הוא שתי הטענות הבאות:

- לכל פרמוטציה σ שאינה הזהות, $k(\sigma) \geq n - 2$. הסיבה לכך היא שאם נבחר צומת u כך ש־$\sigma(u) \neq u$, אז לכל v השונה מ־u ומ־$\sigma(u)$ יתקיים $\{u, v\} \neq \{\sigma(u), \sigma(v)\}$ (כן יתכן אבל שהזוג $\{u, \sigma(u)\}$ "מתהפך" במקום לעבור לזוג שונה).
- לכל פרמוטציה σ המעבירה לפחות $\sqrt{n}$ צמתים ממקומם, $k(\sigma) \geq n^{3/2}/2$. הסיבה לכך היא שאם ניקח צמתים $u_1, \ldots, u_{\sqrt{n}}$ המועברים ממקומם, וצמתים $v_1, \ldots, v_{n/2}$ כך שאף v_j לא שווה לאף u_i ולאף $\sigma(u_i)$, אז כל הזוגות האפשריים $\{u_i, v_j\}$ מקיימים $\{u_i, v_j\} \neq \{\sigma(u_i), \sigma(u_j)\}$.

עתה לא נותר אלא לסכם את ההסתברויות: מספר הפרמוטציות אשר מעבירות פחות מ־$\sqrt{n}$ צמתים ממקומם חסום ע"י $\binom{n}{\sqrt{n}}(\sqrt{n})! \leq 2^{\sqrt{n}\log n}$, ולכן הסיכוי שאחת או יותר מתוכן תהיה אוטומורפיזם חסום לפי איחוד מאורעות על ידי $2^{\sqrt{n}\log n-(n-2)/2} = o(1)$. מספר כל שאר הפרמוטציות חסום ע"י $n! \leq 2^{n\log n}$, והסיכוי שאחת או יותר תהיה אוטומורפיזם חסום ע"י $2^{n\log n - n^{3/2}/4} = o(1)$. לסיכום קיבלנו שאיחוד כל המאורעות E_σ מתקיים בהסתברות $o(1)$, כמבוקש.

הכל קשור

אנחנו נשתמש בקריטריון שגרף הוא קשיר אם אין חלוקה של קבוצת הצמתים לשתי קבוצות לא־ריקות ללא קשתות ביניהן. באופן מדויק, עבור מרחב ההסתברות $G(n, \alpha\ln(n)/n)$ כאשר נסמן את קבוצת הצמתים ב־$V = [n]$, נראה שההסתברות לקיום $1 \leq k \leq n/2$ וקבוצה $U \subset V$ מגודל k ללא קשתות בין U ל־$V \setminus U$ שואפת לאפס כאשר n שואף לאינסוף.

עבור k ו־U ספציפים, ההסתברות שאין קשתות בין U ל־$V \setminus U$ היא $(1-\frac{\alpha\ln(n)}{n})^{k(n-k)} < e^{-\alpha k(n-k)\ln(n)/n}$. כמו כן, מספר הקבוצות האפשריות $U \subset V$ מגודל k הוא $\binom{n}{k} \leq (\frac{en}{k})^k$. ההסתברות שעבור k קבוע קיימת קבוצה U כזו חסומה אם כך ע"י $(\frac{en}{k})^k \cdot e^{-\alpha k(n-k)\ln(n)/n}$. מכאן צריך לפצל למקרים לפי הגודל של k. עבור המשך הניתוח, נבחר β שרירותי שמקיים $1 < \beta < \min\{\alpha, \frac{3}{2}\}$.

עבור $1 \leq k \leq 2/(\beta-1)$ מתקיים $(\frac{e}{k})^k \leq 3$, וכן עבור n גדול דיו מתקיים $\alpha(n-k)/n > \beta$, ואז הביטוי של חסם הסיכוי הוא לכל היותר $3n^k e^{-\beta k\ln(n)} = 3n^{(1-\beta)k} = o(1)$.

עתה נחשב חסם עבור $2/(\beta-1) < k \leq n^{2/3}$ (החזקה של n היא שרירותית, כל ערך גדול מ־$\frac{1}{2}$ וקטן מ־1 היה עובד כאן). שוב עבור n גדול מספיק מתקיים $\alpha(n-k)/n > \beta$, וכן מתקיים $(\frac{e}{k})^k < 1$. על כן חסם הסיכוי כאן הוא לכל היותר $n^{(1-\beta)k} < n^{-2} = o(1/n)$.

לבסוף ננתח את $n^{2/3} < k \leq n/2$. כאן נשתמש ב־$k(n-k)/n \geq \frac{1}{2}k$ וגם ב־$k^k > e^{(2/3)k\ln(n)}$, ונקבל עבור n גדול דיו $(\frac{en}{k})^k \cdot e^{-\alpha k(n-k)\ln(n)/n} < \exp(k + (1-2/3-\alpha/2)k\ln(n)) < e^{-k\ln(n)/6} = o(1/n)$.

עתה נעשה איחוד מאורעות נוסף ונסכום את החסמים על כל ה־k־ים האפשריים. יש בסכום $\lfloor 2/(\beta-1)\rfloor = O(1)$ איברים שהוכחנו שהם $o(1)$, ועוד $\lfloor n/2\rfloor - \lceil 2/(1-\beta)\rceil = O(n)$ איברים שהוכחנו שהם $o(1/n)$. סך כל חסמי ההסתברות המתקבל כך הוא $o(1)$, שזה מה שרצינו להוכיח.

הערות: אם היינו מסתפקים בהוכחת טענת השאלה עבור $\alpha > 2$, הניתוח נהיה קל בהרבה ולא צריך כזה פיצול למקרים (משתמשים ב־$(n-k)/n \geq \frac{1}{2}$ בכל המקרים). עבור $\alpha < 1$ ניתן להוכיח שבהסתברות שואפת ל־1 יהיה צומת מבודד ובפרט הגרף לא יהיה קשיר. מה שקורה עבור $\alpha = 1$ מטופל בשאלות "זאב בודד" ו"למתוח את הקשר" בהמשך החוברת. לבסוף, שימו לב שהחסמים שהפתרון כאן משתמש בהם הם די "רופפים" פרט למקרים של ערכים קטנים של k, ובאמת "המכשול המרכזי" לקשירות הוא רכיבים מגודל קבוע. אפשר למצוא הוכחה פורמלית של האינטואיציה הזו וכן חסמים יותר מדויקים על ההסתברות הדרושה להבטחת קשירות בספר Random Graphs של Bollobás.

פגיעות במרחב

אם A אינה פורשת את כל המרחב הלינארי $V = (\mathbb{Z}_2)^n$, אז תת־המרחב הנפרש $\mathrm{Span}(A)$ הוא ממימד $n-1$ לכל היותר. נסמן ב־$\mathcal{B} = \{V \setminus U : \dim(U) = n-1\}$ את משפחת כל המשלימים של תתי־מרחב $U \subset V$ ממימד $n-1$ בדיוק. אם מתקיים $\mathrm{Span}(A) \neq V$ אז קיים $B \in \mathcal{B}$ שעבורו $\mathrm{Span}(A) \cap B = \emptyset$ ולכן גם $A \cap B = \emptyset$ (יכול להיות יותר מ־B אחד כזה, אם $\dim(\mathrm{Span}(A)) \leq n-2$).

עתה נספור את $|\mathcal{B}|$. כל תת־מרחב $U \subset V$ ממימד $n-1$ יכול להיות מוגדר לפי וקטור $v \in V$ כך שלכל $u \in U$ מתקיים $u \cdot v = \sum_{i=1}^{n} u_i v_i = 0$ (זכרו שהסכום הוא מעל $\mathbb{Z}_2$). למציאת v לוקחים בסיס של U, ומקבלים מערכת הומוגנית ולא מנוונת של $n-1$ משוואות לינאריות ב־n משתנים. הפתרון עבור v הוא יחיד עד כדי כפל בסקלר שונה מ־0, רק שב־$\mathbb{Z}_2$ אין איבר שונה מ־0 למעט 1, כך ש־v הוא יחיד. אם נשים לב לכך שההעתקה מ־U ל־v היא הפיכה (בהינתן v אפשר "לשחזר" את $U = \{v \in V : u \cdot v = 0\}$), נקבל שמתקיים $|\mathcal{B}| = |V \setminus \{0\}| = 2^n - 1$. נשים לב גם שמתקיים $|B| = 2^n - 2^{n-1} = 2^{n-1}$ לכל $B \in \mathcal{B}$, כי הגודל של מרחב ממימד k מעל $\mathbb{Z}_2$ הוא 2^k.

עבור קבוצה $B \in \mathcal{B}$ בודדת, ההסתברות ש־A אינה מכילה איבר מ־B (לפי מרחב ההסתברות שלנו) היא $(1-2n/2^n)^{|B|} = (1-2n/2^n)^{2^{n-1}} < e^{-n} = o(2^{-n})$. על כן, לפי איחוד מאורעות, ההסתברות שיש קבוצה $B \in \mathcal{B}$ שעבורה מתקיים $A \cap B = \emptyset$ היא $o(1)$, ז"א שבהסתברות $1 - o(1)$ אין קבוצה כזו ואז מתקיים $\mathrm{Span}(A) = V$.

הולכים במעגלים

ראשית נניח שמתקיים $k < n/10$. אחרת חישוב פשוט מראה שיש לפחות $n/20$ מעגלים מגודל שאינו עולה על 20 (כי סכום אורכי כל המעגלים הוא n), ואז אפשר פשוט להגריל באופן יוניפורמי קבוצה של 40 אינדכסים ולבדוק לכל אחד מהם האם הוא במעגל באורך 20 או פחות (ע"י חישוב הערכים $\sigma(v), \sigma(\sigma(v)), \ldots$, סה"כ 20 איטרציות), וסה"כ יש לנו מספר קבוע של קריאות של ערכי σ.

מכיוון שסכום גודלי כל המעגלים בפירוק (כולל "מעגלים מאורך 1") הוא בדיוק n, אם יש לפחות k מעגלים, אז חישוב פשוט אומר שיש לפחות $k/2$ מעגלים שגודלם אינו עולה על $2n/k$. עבור $0 \leq j \leq \log(2n/k)$, נסמן ב־m_j את מספר המעגלים שגודלם בין 2^j ל־$2^{j+1}-1$. סכום ה־m_j הנ"ל הוא לפחות $k/2$, ומצד שני יש לא יותר מ־$\lfloor \log(2n/k) \rfloor + 1 \leq 2\log(n/k)$ ערכים אפשריים עבור j. לכן קיים j_* שעבורו $m_{j_*} \geq k/4\log(n/k)$.

האלגוריתם יעבוד כך: עבור כל $0 \leq j \leq \log(2n/k)$, נבחר באופן מקרי, יוניפורמי וב"ת $10\log(n/k)n/k2^j$ אינדכסים (עם חזרות, ז"א שמרשים שאותו אינדכס ייבחר יותר מפעם אחת). לכל אינדכס v שנבחר עבור ה־j הנ"ל, נבדוק האם הוא במעגל מגודל קטן מ־2^{j+1} ע"י קריאת $\sigma(v), \sigma(\sigma(v)), \sigma^3(v), \ldots, \sigma^{2^{j+1}}(v)$. סה"כ, עבור כל ערך נתון של j אנחנו מבצעים $10 \cdot 2^{j+1}\log(n/k)n/k2^j = O((n/k)\log(n/k))$ קריאות של ערכי σ, וסה"כ לכל ערכי j שאנחנו בודקים אנחנו מבצעים $O((n/k)(\log(n/k))^2)$ קריאות של ערכי σ.

על מנת לסיים צריך להראות שהאלגוריתם מוצא מעגל בהסתברות לפחות $\frac{1}{2}$. אנחנו נראה שזה קורה בסבב הבדיקה של $j = j_*$ (האלגוריתם בודק את כל ערכי j האפשריים כי הוא לא יודע את j_* מראש). בשלב זה, הסיכוי שלא בחרנו אף אינדכס שנמצא במעגל שגודלו בין 2^{j_*} ל־$2^{j_*+1}-1$ חסום ע"י $\frac{1}{2} > (1-2^{j_*}m_{j_*}/n)^{10\log(n/k)n/k2^{j_*}}$. הסבר לביטוי: מספר האינדכסים הכולל במעגלים עם הגדלים הנ"ל הוא לפחות $2^{j_*}m_{j_*}$, ולכן הסיכוי לא למצוא

אינדכס כזה בכל איטרציה חסום ע"י $(1-2^{j_*}m_{j_*}/n)$. מכיוון שבצענו $10\log(n/k)n/k2^{j_*}$ איטרציות באופן ב"ת, חסם ההסתברות הכולל הוא $(1-2^{j_*}m_{j_*}/n)^{10\log(n/k)n/k2^{j_*}}$. בהצבת החסם על m_{j_*} תוך שימוש באי השוויון $(1-x)\le e^{-x}$ מתקבל המבוקש. ברגע שמצאנו לפחות אינדכס אחד כזה, אנחנו אכן נגלה שגודל המעגל המכיל אותו קטן מ-2^{j_*+1}, כי את הבדיקה הזו עשינו באופן דטרמיניסטי לכל אחד מהאינדכסים שנבחרו.

קשה לכסות

נגריל באופן יוניפורמי וב"ת תתי־קבוצה $B^{(j)}\subseteq\{1,\dots,n\}$ תחת התנאי ש-$|B^{(j)}|$ הוא אי־זוגי (לכל תת־קבוצה יש 2^{n-1} אפשרויות, וסה"כ יש $2^{(n-1)k}$ אפשרויות לסדרה $B^{(1)},\dots,B^{(k)}$, שכל אחת מהן מתקיימת בהסתברות שווה). נסמן את המכפלה הקרטזית $B=\prod_{j=1}^k B^{(k)}\subseteq\{1,\dots,n\}^k$. נשים לב שבפרט $|B|=\prod_{j=1}^k|B^{(k)}|$ הוא אי־זוגי תמיד. אם $\bigcup_{i=1}^m A_i=\{1,\dots,n\}^k$, אז $|B\cap\bigcup_{i=1}^m A_i|=|B|$ הוא אי־זוגי תמיד. מצד שני תמיד מתקיים $|B\cap\bigcup_{i=1}^m A_i|=\sum_{i=1}^m|B\cap A_i|=\sum_{i=1}^m\prod_{j=1}^k|B^{(j)}\cap A_i^{(j)}|$, כי הקבוצות $A_1,\dots,A_m$ הן זרות לפי נתוני השאלה.

נראה עבור i קבוע שמתקיים ש-$\prod_{j=1}^k|B^{(j)}\cap A_i^{(j)}|$ יהיה אי־זוגי בהסתברות 2^{-k}. מכך נובע שאם $m<2^k$ אז בהתסברות חיובית כל הערכים $|B\cap A_i|$ הם זוגיים, ולכן גם $|B\cap\bigcup_{i=1}^m A_i|$ זוגי, אבל כזכור זה אינו יכול לקרות בהסתברות חיובית אם מתקיים $\bigcup_{i=1}^m A_i=\{1,\dots,n\}^k$.

לשם כך נשים לב שכאשר $A_i^{(j)}$ אינו שווה ל-$\{1,\dots,n\}$ (כפי שאכן נתון בשאלה), אז $|B^{(j)}\cap A_i^{(j)}|$ יהיה אי־זוגי בהסתברות $\frac{1}{2}$ בדיוק (החיתוך יהיה תת־קבוצה מקרית יוניפורמית של $A_i^{(j)}$; אנחנו מתעלמים מהמקרה שהקבוצה A_i ריקה כי את אלו אפשר פשוט למחוק מהרשימה). מכיוון שה-$B^{(j)}$ נבחרות באופן ב"ת, זה אומר שהסיכוי ש-$|B^{(j)}\cap A_i^{(j)}|$ יהיה אי־זוגי לכל ה-j הוא 2^{-k}. על מנת שהמכפלה $\prod_{j=1}^k|B^{(j)}\cap A_i^{(j)}|$ תהיה אי־זוגית צריך שכל האיברים יהיו אי־זוגיים, ז"א שהסיכוי לכך הוא גם 2^{-k}.

גרף מקרי ויחיד

ראשית נוכיח כי עבור גרף מקרי כזה התנאי הבא מתקיים בהסתברות 1: לכל קבוצת צמתים סופית U ולכל $U'\subseteq U$, קיים צומת כל שהוא $v\notin U$ כך ש-U' כולם שכניו ו-$U\setminus U'$ כולם לא־שכניו.

בהינתן קבוצה סופית מסויימת U מגודל k, תת קבוצה $U'\subseteq U$ וצומת $v\notin U$, הסיכוי ש-v יהיה מחובר לכל צמתי U' ואינו מחובר לכל צמתי $U\setminus U'$ הוא בדיוק 2^{-k}. נסמן מאורע זה ב-A_v. קבוצת המאורעות $\{A_v : v\notin U\}$ היא בלתי תלויה לחלוטין, ולכן הסיכוי שאף אחד מהם לא יקרה הוא 0 (כגבול של $(1-2^{-k})^l$ עבור $l\to\infty$). נסמן ב-$B_{U,U'}$ את המאורע שאין צומת v כנדרש. ההסתברות למאורע זו היא 0, ומכיוון שיש מספר בן מניה של מאורעות $B_{U,U'}$ אפשריים, הסיכוי שיש אילו שהם U,U' ללא v מתאים הוא 0 (החסם על איחוד מאורעות עובד כל עוד מספרם הוא בן מניה).

הערה: במרחבי הסתברות אינסופיים יש הבדל בין "מאורע בהסתברות 0" לבין "מצב שאינו אפשרי". אפשר לתאר גרפים אינסופיים עבורם לא לכל U,U' יש v מתאים, אבל ההסתברות שיתקבל גרף דווקא מקבוצה זו היא 0.

עתה נניח ש-G ו-G' הם שני גרפים שנבחרו לפי $G(\mathbb{N},\frac{1}{2})$. בהסתברות 1, לכל U סופית ו-$U'\subseteq U$ יש צומת $v\notin U$ כך שב-G יש קשתות ממנו ל-U' ולא ל-$U\setminus U'$, וצומת $v'\notin U$ כך שב-G' יש קשתות ממנה ל-U' ולא ל-$U\setminus U'$. נסמן את הצמתים הנ"ל ב-$v_{U,U'}$ ו-$v'_{U,U'}$ בהתאמה (אם יש יותר מאפשרות אחת, נבחר את זו עם האינדכס הנמוך ביותר).

עתה נבנה באינדוקציה קבוצות W_i ו-W'_i ופונקציות חח"ע ועל $f_i:W_i\to W'_i$, כך שיתקיימו הדברים הבאים:

- לכל $i<j$ מתקיים $W_i\subseteq W_j$, $W'_j\subseteq W'_j$, ו-$f_j|_{W_i}=f_i$ (ז"א ש-f_j היא הרחבה של f_i).

- לכל i הפונקציה f_i היא איזומורפיזם מהגרף המושרה ע"י G על W_i אל הגרף המושרה ע"י G' על W'_i (ז"א שלכל $u, v \in W_i$, uv היא קשת של G אם ורק אם $f_i(u)f_i(v)$ היא קשת של G').
- מתקיים $\bigcup_{i\in\mathbb{N}} W_i = \bigcup_{i\in\mathbb{N}} W'_i = \mathbb{N}$.

לאחר הבניה הנ"ל ניתן לבנות את האיזומורפיזם $f : \mathbb{N} \to \mathbb{N}$ ע"י כך שלכל $v \in \mathbb{N}$ נבחר את $f(v)$ להיות שווה ל־$f_i(v)$ עבור i המקיים $v \in W_i$. הפונקציה f מוגדרת על כל $\mathbb{N}$ בגלל הסעיף השלישי למעלה, ומוגדרת היטב בגלל הסעיף הראשון למעלה. היא חח"ע כי לכל $u \neq v$ ניתן לבחור i כך ששניהם ב־W_i ולבדוק את f_i, היא על כי לכל w ניתן לבחור i כך ש־$w \in W'_i$ (קיים לפי הסעיף השלישי) ולמצוא את $f^{-1}(w) = f_i^{-1}(w)$, והיא איזומורפיזם בין גרפים לפי הסעיף השני למעלה.

נותר אם כן לבנות את הקבוצות W_i ו־W'_i ואת הפונקציות f_i. בסיס האינדוקציה יהיה $W_0 = W'_0 = \emptyset$, כאשר f_0 היא ה"פונקציה" הטריביאלית ביניהן. נניח שבנינו את הקבוצות והפונקציה עבור i, ונראה את הבניה עבור $i+1$. אנו נפצל למקרים לפי הזוגיות של i. כאן נשתמש במוסכמה שהמספרים הטבעיים מתחילים מ־0.

עבור $i = 2k$, אם $k \in W_k$ אז פשוט נגדיר $W_{k+1} = W_k$, $W'_{k+1} = W'_k$ ו־$f_{k+1} = f_k$. אחרת, ראשית נסמן ב־U_k את קבוצת השכנים של k ב־W_k, וב־$U'_k \subseteq W'_k$ את התמונה שלהם לפי f_k. נגדיר עתה $W_{k+1} = W_k \cup \{k\}$ ו־$W'_{k+1} = W'_k \cup \{v'_{W'_k,U'_k}\}$. נגדיר $f_{k+1}(k) = v'_{W'_k,U'_k}$, ובשאר המקומות f_{k+1} תהיה זהה ל־f_k. הנחת האינדוקציה והידוע על $v'_{W'_k,U'_k}$ (כולל זה שאינו ב־W'_k) מבטיחה שהבניה תתן את הקבוצות והפונקציה המבוקשות.

עבור $i = 2k+1$, אם $k \in W'_k$ אז פשוט נגדיר $W_{k+1} = W_k$, $W'_{k+1} = W'_k$ ו־$f_{k+1} = f_k$. אחרת, ראשית נסמן ב־U'_k את קבוצת השכנים של k ב־W'_k, וב־$U_k \subseteq W_k$ את קבוצת המקורות שלהם לפי f_k (זכרו שזוהי פונקציה חח"ע ועל). נגדיר עתה $W'_{k+1} = W'_k \cup \{k\}$ ו־$W_{k+1} = W_k \cup \{v_{W_k,U_k}\}$. נגדיר $f_{k+1}(v_{W_k,U_k}) = k$, ובשאר המקומות f_{k+1} תהיה זהה ל־f_k. הנחת האינדוקציה והידוע על v_{W_k,U_k} (כולל זה שאינו ב־W_k) מבטיחה שהבניה גם כאן תתן את הקבוצות והפונקציה המבוקשות.

כל התנאים המבוקשים פרט לדרישת האיחוד בסעיף השלישי יתקיימו באינדוקציה, וכן לכל k מובטח שהוא נמצא ב־W_{2k+1} וב־W'_{2k+2}, וכך מתקיימת גם דרישת האיחוד. בזאת סיימנו את המבוקש.

לינאריות התוחלת

השרשרת

עבור ערכים $p_1, \ldots, p_n \in [0,1]$ נתונים, נגדיר מרחב הסתברות על בחירה של קבוצה $S \subseteq \{0,1\}^n$ באופן הבא: ראשית בהסתברות $\frac{1}{2}$ בדיוק נקבע שמתקיים $1 \in S$ ובהסתברות $\frac{1}{2}$ נקבע שלא מתקיים $1 \in S$. נמשך מכאן באינדוקציה על i: לאחר שקבענו את $S \cap \{1, \ldots, i-1\}$, נחלק למקרים לפי האם בחרנו שמתקיים $i-1 \in S$. אם הדבר מתקיים, נקבע $i \in S$ בהסתברות p_i ונקבע $i \notin S$ בהסתברות $1-p_i$, כאשר הבחירה החדשה נעשית באופן ב"ת בבחירות אקראיות קודמות. אם לא מתקיים $i-1 \in S$, אז נקבע $i \in S$ בהסתברות $1-p_i$ ונקבע $i \notin S$ בהסתברות p_i. הראו שמתקיים $\mathrm{E}[|S|] = \frac{n}{2}$.

הלינאריות הלא־נכונה

תנו דוגמה למרחב הסתברות המקיים את הדברים הבאים עבור המשתנים המקריים Y ו־$X_1, X_2, \ldots$ (סדרה אין־סופית בת מניה של משתנים):

- המשתנה Y מתפלג מעל המספרים הטבעיים, התוחלת שלו סופית, אבל הוא יכול לקבל כל מספר טבעי בהסתברות גדולה מ־0.
- המשתנים X_i כולם ב"ת זה בזה (אבל יכולים להיות תלויים ב־Y) ומתפלגים מעל $\{0,1\}$.

- אם נגדיר את המ"מ $Z = \sum_{j=1}^{Y} X_j$, **לא קיים** אף $k \in \mathbb{N}$ שעבורו מתקיים $\mathrm{E}[Z|Y=k] = \sum_{j=1}^{k} \mathrm{E}[X_j]$.

הערה: ה"שוויון" שבסעיף השלישי מאוד מפתה, ולא אחת יוצא להיתקל בטיעון (שגוי כמובן) ש"משתמש" בו. מטרת התרגיל היא להראות שהוא אינו מובטח (השוויון $\mathrm{E}[Z|Y=k] = \sum_{j=1}^{k} \mathrm{E}[X_j|Y=k]$ ש"דומה" לשוויון הנ"ל כן מתקיים, כמקרה פרטי של לינאריות התוחלת במרחב ההסתברות המותנה המתאים).

גרפים רחוקים

המרחק בין שני גרפים G ו־H מוגדר כמספר הקשתות המינימלי שיש להוריד ו/או להוסיף ל־H כך שיהפוך להיות גרף איזומורפי ל־G. הראו שאם ל־G יש n צמתים ו־$m \neq \frac{1}{2}\binom{n}{2}$ קשתות אז הגרף הרחוק ביותר מ־G הוא או הגרף המלא או הגרף הריק.

רמז: אפשר להראות לכל H שהמרחק שלו מ־G חסום ע"י $p(\binom{n}{2} - m) + (1-p)m$ עבור $0 \le p \le 1$ מתאים.

הגעה מהוססת

מגרילים משתנים מקריים $X_1, X_2, \ldots$, כולם בלתי תלויים ויוניפורמים מתוך $\{0,1\}$. נסמן ב־T_k את המספר הקטן ביותר עבורו מתקיים $\sum_{t=1}^{T_k} X_t \ge k$. במילים אחרות, זהו ה"זמן" שבו מגיעים למספר k אחרי שבכל שלב מחליטים בהסתברות $\frac{1}{2}$ אם נשארים במקום או עולים ב־1. בפרט, $T_0 = 0$ תמיד ו־T_1 מתפלג כמו ההתפלגות הגיאומטרית עם פרמטר $\frac{1}{2}$. חשבו את התוחלת $\mathrm{E}[T_k]$.

קשתות לטווח קצר

נתון גרף דו־צדדי $G = (U, V, E)$, כאשר קבוצות הצמתים U ו־V שתיהן בגודל n, והקבוצה E בת αn^2 קשתות. עבור צומת $u \in U$ נסמן ב־$F_u \subseteq E$ את קבוצת הקשתות הנוגעות בשכנים של u (שימו לב שבפרט זה כולל גם קשתות מ־u עצמו). הראו שקיים u עבורו $|F_u| \ge \alpha^2 n^2$.

גלגלים מסתובבים

עבור $r \in \mathbb{N}$ נתון, תהי V קבוצה של kr נקודות על מעגל, מרווחות במרווחים שווים. אפשר להניח שהמדובר במעגל היחידה ושמתחילים מהנקודה על ציר ה־x החיובי, ואז, למשל עבור $k = r = 2$, המדובר בקבוצה $\{(1,0), (0,1), (-1,0), (0,-1)\}$. יהיו $V_1, \ldots, V_r$ קבוצות המחלקות את V באופן שווה (ז"א שכל V_i היא ת"ק של V מגודל k, וקבוצות אלו זרות זו לזו). הראו שאם k גדול מספיק (כפונקציה של r) אז אפשר למצוא בכל V_i קודקודים של משולש $t_i = \{u_i, v_i, w_i\}$ (לא מנוון, הקודקודים שונים זה מזה) כך ש־$t_1, \ldots, t_r$ חופפים.

הערה: לא צריך להתעסק בגיאומטריה של המישור, אפשר להתייחס לזה כשאלה בחישוב בשלמים מודולו kr.

קירבה לדרגה קבועה

נתבונן בגרף המקרי $G(n,p)$ עבור $p = \alpha/n$ (כזכור, המדובר בגרף עם קבוצה V בעלת n צמתים, כך שכל זוג צמתים uv נבחר להיות בקבוצת הקשתות E בהסתברות p בדיוק, באופן ב"ת לחלוטין בבחירות של הזוגות האחרים). כמו כן נתונים β ו־γ, כולם גדולים מ־0. הראו את קיומו של d שתלוי ב־α, β, γ בלבד, כך שבסיכוי לפחות $1 - \gamma$ אפשר להסיר מהגרף עד βn קשתות, ולגרף שיוותר תהיה דרגה מקסימלית שאינה עולה על d.

סיפוק חלקי של נוסחת CNF

נתונה נוסחת CNF עם m פסוקיות, כל אחת מהן דיסיונקציה ("או") בין כמה ליטרלים (משתנים או שלילתם). מובטח כי לא קיימות פסוקיות ריקות, וכי לא קיים משתנה x_i עבורו מופיעות שתי הפסוקיות $x_i, \neg x_i$ (כלומר עבור כל שתי פסוקיות נתונות, ניתן לספק את שתיהן בו זמנית). הוכיחו כי קיימת השמה שמספקת לפחות

$\alpha = \frac{1}{2}\left(\sqrt{5}-1\right)$ מהפסוקיות. הערה: תרגיל זה הוא משפט של Specker, אך ההוכחה המקורית אינה משתמשת בשיטה ההסתברותית.

רמז: כדאי לחשוב על המקרה שבו כל הפסוקיות הן מהצורה "x_i" או "$\neg x_j \vee \neg x_k$".

זאב בודד

עבור הגרף המקרי $G(n, \ln(n)/n)$, הראו שבהסתברות $\Omega(1)$ יהיה לפחות צומת אחד ללא שכנים (ובפרט הגרף לא יהיה קשיר).

הערה ורמז: משמעות השאלה היא שיש להראות חסם תחתון קבוע (לא תלוי n) גדול מאפס על ההסתברות הנ"ל, שתקף לכל n גדול דיו. עוזר לפתח ביטוי עבור תוחלת מספר הצמתים חסרי השכנים. תוחלת זו שואפת ל־1, וניתן להראות שמרבית ה"תרומה" לתוחלת לא מגיעה ממקרים עם מספר רב של צמתים כאלו.

רב־קרב תרנגולים

נתונים n תרנגולים במעגל (אפשר להניח ש־n גדול דיו, מספיק שמתקיים $n \geq 5$). בהתחלה כל התרנגולים בריאים, אבל בכל סיבוב כל תרנגול בריא בוחר אם לנקר את התרנגול מימינו או משמאלו באופן יוניפורמי וב"ת באחרים (יכול להיות ששני תרנגולים בריאים ינקרו זה את זה בו זמנית). תרנגול שכבר ניקרו אותו אינו מנקר יותר בסיבובים הבאים, אבל יכול להיות שתרנגול לידו ינקר אותו שוב.

התהליך מפסיק כשאין יותר שני תרנגולים בריאים (לא מנוקרים) זה ליד זה. הראו שתוחלת מספר התרנגולים הבריאים בסוף התהליך היא $n/6$.

הדרכה: כדאי להגדיר תרנגול כ"בטוח" אם הוא בריא, וכבר אין לידו תרנגול בריא. כדאי בתור שלב ראשון לחשב את תוחלת מספר התרנגולים הבטוחים שנשארו לאחר סיבוב הניקורים הראשון ואת תוחלת מספר התרנגולים הבריאים הלא־בטוחים שנשארו אז. עבור הסיבובים הבאים תוצאת השאלה "דו־קרב תרנגולים" (מהפרק על בניה וניתוח של מרחבי הסתברות) יכולה לעזור.

איזונים ובלמים

עבור קבוצה סופית לא ריקה S_0 ופונקציה $f : S_0 \to \{0,1\}$ נבצע את התהליך הבא: בשלב ה־i (כאשר מתחילים מ־$i=0$), אם יש ל־$f|_{S_i}$ רק ערכי 0 אז נפלוט "0", אם יש ל־$f|_{S_i}$ רק ערכי 1 אז נפלוט "1", ובכל מקרה אחר נבחר קבוצה S_{i+1} באופן יוניפורמי מבין כל תתי־הקבוצה של S_i למעט $\emptyset$ ו־S_i עצמה, ונמשיך לשלב הבא. הראו שההסתברות לקבל פלט "1" בסוף התהליך היא בדיוק $|T_0|/|S_0|$, כאשר T_0 היא הקבוצה $\{s \in S_0 : f(s) = 1\}$ (במילים אחרות, ההסתברות לקבל "1" בסוף התהליך שווה לכמות היחסית של ערכי ה־1 בפונקציה f).

הדרכה: לכל i שבו האלגוריתם מתבצע, נסמן ב־T_i את הקבוצה $\{s \in S_i : f(s) = 1\}$. ראשית דבר הוכיחו שמתקיים $\mathrm{E}[|T_{i+1}|/|S_{i+1}|] = \mathrm{E}[|T_i|/|S_i|]$.

לבלוע את החוכמה

נתונה משפחה $\mathcal{F}$ של תתי־קבוצה של $\{1,\ldots,n\}$. עבור $0 \leq q \leq n$ ועבור $0 < p \leq 1$, נגדיר שני מרחבי הסתברות להגרלת תת־קבוצה $Q \subseteq \{1,\ldots,n\}$.

- במרחב μ, נגריל את Q באופן יוניפורמי מכל $\binom{n}{q}$ האפשרויות לתת־קבוצה בגודל q של $\{1,\ldots,n\}$.
- במרחב ν, נבחר כל אינדכס $1 \leq i \leq n$ להיות איבר ב־Q בהסתברות p, באופן ב"ת לחלוטין בתוצאות ההגרלה עבור אינדכסים אחרים.

נסמן ב־A את המאורע "קיימת $F \in \mathcal{F}$ שעבורה מתקיים $F \subseteq Q$". הראו שמתקיים $\Pr_\mu[A] > \Pr_\nu[A] - \frac{pn}{q}$.

הדרכה: זוהי בעיקר שאלה על בניה וניתוח של מרחבי הסתברות (החלק שקשור בלינאריות התוחלת אינו קשה). נסו להגדיר מרחב הסתברות על זוגות (Q, Q') כך ש־Q מתפלגת לפי ν ו־Q' מתפלגת לפי התפלגות עם מרחק לא גדול מדי מ־μ.

כל אחד את עצמו

נתון גרף (פשוט, לא מכוון) G עם n צמתים ודרגה חסומה מלמעלה ע"י d. ננסה לצבוע את הגרף ע"י $3d$ צבעים באופן הבא:

- לפני השלב הראשון כל הצמתים צבועים בצבע "1".
- בכל שלב, כל צומת שיש לו שכן שצבוע באותו צבע כמוהו, בוחר באופן יוניפורמי (וב"ת בשלבים קודמים ו/או צמתים אחרים) צבע מתוך הקבוצה $\{1, \ldots, 3d\}$ (בפרט יש סיכוי קטן שהוא ישאר באותו צבע).

הראו שתוחלת מספר השלבים עד שמצאנו צביעה חוקית (לאף צומת אין שכן מאותו צבע) היא $O(\log(n))$.

רמז: כדאי קודם לחסום את תוחלת מספר הצמתים שיש להם שכן מאותו צבע בשלב ה־k.

הערה: אלו מכם שלמדו חישוב מבוזר בוודאי מזהים את צורת הצגת האלגוריתם. זהו מודל חישובי שבו יש מעבד בכל צומת של הגרף, וכל התקשורת בין המעבדים מתבצעת רק דרך קשתות הגרף.

לתפוס את המרובה

עבור מרחב התפלגות μ מעל $\mathbb{N}$, נגדיר מ"מ $X_1, \ldots, X_k$ שכולם ב"ת, וכל אחד מהם לחוד מתפלג לפי μ. נסמן ב־A את קבוצת הערכים המופיעים ב־$X_1, \ldots, X_k$ (קבוצה רגילה, לא מחשיבים כפילויות). הראו שקיים קבוע גלובלי C שעבורו אם מתקיים $k \geq Cn/\epsilon$ (לכל $n \in \mathbb{N}$ ולכל $\epsilon > 0$), אז בהסתברות לפחות $\frac{1}{2}$ מתקיים איחוד המאורעות "$\mu(A) \geq 1 - \epsilon$" ו־"$|A| \geq n$" (הבהרה – "איחוד" פירושו שלפחות אחד משני המאורעות מתקיים).

הערה לפורמליסטים: סדרת המשתנים $X_1, \ldots, X_k$, ויחד אתה המשתנים המקריים שערכם מוגדר להיות $\mu(A)$ ו־$|A|$ בהתאמה למעלה, יכולים להיות מוגדרים מעל מרחב התפלגות מתאים שקבוצת הבסיס שלו היא $\mathbb{N}^k$. המרחב הנ"ל (למרות שאינו סופי) הוא בדיד, ובפרט ההסתברויות לכל המאורעות נקבעות ע"י ההסתברויות לאיברים של קבוצת הבסיס.

רמז: כדאי לחשוב על סדרה לא מוגבלת מראש $X_1, X_2, \ldots$ של מ"מ, ולנתח את ה־k הקטן ביותר שעבורו קבוצת הערכים של $X_1, \ldots, X_k$ מקיימת את הנדרש.

לא כל הדרכים מובילות

הילוך מקרי מצומת s בגרף G (שיכול להיות מכוון) מוגדר באופן הבא: המשתנה המקרי X_0 (שמקבל ערכים מתוך קבוצת הצמתים של הגרף) יקבל בהסתברות 1 את הצומת s. המ"מ X_1 יקבל שכן של s שנבחר באופן יוניפורמי מקבוצת השכנים האפשריים. בהמשך, לאחר בחירת ערך X_{i-1}, המ"מ X_i יקבל צומת שנבחר יוניפורמית מקבוצת השכנים של הערך של X_{i-1}, כאשר כל הגרלה נערכת באופן ב"ת בהגרלות קודמות.

נסמן ב־T_v את זמן ההגעה הראשון לצומת v, ז"א את המספר הכי נמוך עבורו $X_{T_v} = v$. למשל, ברור שמתקיים $X_s = 0$ בהסתברות 1. הראו (ללא שימוש בכלים מתקדמים של ניתוח הילוכים מיקריים) שלכל גרף G וצומת s קיים צומת v עבורו $\mathrm{E}[T_v] = \Omega(n)$, כאשר n מציין את מספר הצמתים בגרף.

פתרונות לתרגילים על לינאריות התוחלת

השרשרת

ראשית נראה באינדוקציה שלכל $1 \leq i \leq n$ מתקיים $\Pr[i \in S] = \frac{1}{2}$ (הסתברות לא מותנה). עבור $i = 1$ הדבר נובע ישירות מתיאור הבניה של S. עתה נניח שעבור $2 \leq i \leq n$ מתקיים $\Pr[i-1 \in S] = \frac{1}{2}$, ונראה מכך את המבוקש עבור i באמצעות נוסחת ההסתברות השלמה:

$$\Pr[i \in S] = \Pr[i \in S | i-1 \in S]\Pr[i-1 \in S] + \Pr[i \in S | i-1 \notin S]\Pr[i-1 \notin S] = \frac{p_i}{2} + \frac{1-p_i}{2} = \frac{1}{2}$$

לאחר שהוכחנו זאת, נגדיר לכל $1 \leq i \leq n$ את מ"מ האינדיקטור X_i שיהיה שווה ל־1 אם $i \in S$ ושווה ל־0 אם $i \notin S$. אלו אינם משתנים ב"ת (אלא אם כן כל ה־p_i שווים ל־$\frac{1}{2}$) אבל הם כן מקיימים $\mathrm{E}[X_i] = \frac{1}{2}$ עבור $1 \leq i \leq n$, ולכן לפי לינאריות התוחלת מתקיים $\mathrm{E}[|S|] = \sum_{i=1}^{n} \mathrm{E}[X_i] = \frac{n}{2}$.

הלינאריות הלא־נכונה

נגדיר ראשית את $X_1, X_2, \ldots$ להיות מ"מ ב"ת לחלוטין שכל אחד מהם מקבל 1 בהסתברות $\frac{2}{3}$ ומקבל 0 בהסתברות $\frac{1}{3}$ (ולכן אלו מקיימים את הסעיף השני בשאלה). עתה נגדיר את Y להיות המ"מ שמקבל את האינדכס הכי קטן i שעבורו $X_i = 1$, ז"א $Y = \min\{i \in \mathbb{N} : X_i = 1\}$. הביטוי הזה מוגדר (המינימום הוא מעל קבוצה לא ריקה) בהסתברות 1. המשתנה Y מקיים $\Pr[Y = k] = \frac{2}{3}(\frac{1}{3})^{k-1}$ (זוהי ההתפלגות הגיאומטרית עם פרמטר $\frac{1}{3}$), ובפרט יש לו הסתברות חיובית לקבל כל מספר טבעי. גם לא קשה לראות שהתוחלת שלו סופית (ושווה ל־$\frac{3}{2}$).

הדבר האחרון לשים לב הוא שבהתניה על $Y = k$ מתקיים $\sum_{i=1}^{k} X_i = 1$ בהסתברות 1, כי אם $Y = k$ אז $X_k = 1$ וגם $X_i = 0$ לכל $i < k$. על כן $\mathrm{E}[Z | Y = k] = 1$, ומצד שני $\sum_{i=1}^{k} \mathrm{E}[X_i] = \frac{2k}{3}$, וזה מספר שאינו שווה ל־1 לאף k.

גרפים רחוקים

נסמן ב־V את קבוצת הצמתים של G וב־V' את קבוצת הצמתים של H כל שהוא (כאשר שתי קבוצות הצמתים מגודל n). נסמן $p = |E'|/\binom{n}{2}$ כאשר E' היא קבוצת הקשתות של H. נבחר עתה פונקציה חח"ע ועל $f : V \to V'$ באופן יוניפורמי (מתוך קבוצת כל הפונקציות הנ"ל), ונחשב את תוחלת מספר הזוגות $u, v \in V$ שיש עבורם הבדל בין השיוך ל־E (קבוצת הקשתות של G) של u, v לבין השיוך ל־E' של $f(u), f(v)$. נשים לב כי הפונקציה f עם הכי מעט הבדלים קובעת את המרחק בין הגרפים.

נסמן ב־$X_{u,v}$ את משתנה האינדיקטור שיקבל 1 אם יש כזה הבדל, ו־0 אחרת. התוחלת שלו היא p אם u, v אינו קשת של G, ו־$1 - p$ אם u, v כן קשת של G. לכן, תוחלת מספר ההבדלים הכולל היא

$$\mathrm{E}[\sum_{u,v} X_{u,v}] = \sum_{uv \notin E} p + \sum_{uv \in E} (1-p) = p(\binom{n}{2} - m) + (1-p)m$$

בפרט זהו חסם עליון על המרחק בין G ל־H, כי קיימת f אחת לפחות שבה מספר ההבדלים אינו עולה על התוחלת, ולכן בפרט בזו האופטימלית מספר ההבדלים חסום על ידי ערך זה.

עתה נשים לב שהפונקציה הנ"ל של p מקבלת את המקסימום שלה עבור $p = 0$ אם $m > \frac{1}{2}\binom{n}{2}$, ומקבלת אותו עבור $p = 1$ אם $m < \frac{1}{2}\binom{n}{2}$ (כזכור הנחנו שלא מתקיים $m = \frac{1}{2}\binom{n}{2}$). במקרה הראשון הגרף הרחוק ביותר היחידי הוא הגרף הריק, שהוא היחידי עבורו $p = 0$ (קל לראות שהמרחק ממנו אכן שווה לחסם במקרה זה), ובמקרה השני הגרף הרחוק ביותר היחידי הוא הגרף המלא, שהוא היחידי עבורו $p = 1$.

הגעה מהוססת

עבור כל k, נחשב ראשית את תוחלת ההפרש $\mathrm{E}[T_k - T_{k-1}]$. זוהי התוחלת של מספר ההטלות של מטבע הוגנת עד שמתקבל "1", אשר כידוע שווה ל־2 (התפלגות ההפרש היא ההתפלגות הגיאומטרית עם פרמטר $\frac{1}{2}$, ואפשר לחשב את התוחלת לפי $\sum_{i=1}^{\infty} i \cdot \Pr[X = i] = \sum_{i=1}^{\infty} \Pr[X \geq i]$). עתה משתמשים בלינאריות התוחלת לחישוב התוחלת שלנו: $\mathrm{E}[T_k] = \mathrm{E}[T_{k-1}] + \mathrm{E}[T_k - T_{k-1}] = \mathrm{E}[T_k] + 2$. מכאן נובע (למשל באינדוקציה) שמתקיים $\mathrm{E}[T_k] = 2k$.

קשתות לטווח קצר

נבחר את u באופן מקרי ויוניפורמי מכל צמתי U, ונחשב את התוחלת של $|F_u|$. לכל צומת $v \in V$ נסמן את הדרגה שלו ב־d_v. עבור קשת wv, הסיכוי שלה להיות ב־F_u הוא בדיוק הסיכוי ש־u ייבחר להיות שכן של v, ז"א d_v/n. על כן, תוחלת מספר הקשתות הסמוכות ל־v שנמצאות ב־F_u היא $(d_v)^2/n$. על מנת לחשב את תוחלת מספר הקשתות הכולל ב־F_u לפי לינאריות התוחלת, נסכום על כל $v \in V$ (כל קשת ב־E סמוכה ל־$v \in V$ אחד בדיוק), ונקבל $\mathrm{E}[|F_u|] = \sum_{v \in V} (d_v)^2/n = n \sum_{v \in V} (d_v/n)^2 \geq (\sum_{v \in V} (d_v/n))^2 = \alpha^2 n^2$.

הסבר: אי השוויון למעלה הוא אי־שוויון הנורמות, ולאחריו השתמשנו ב־$\sum_{v \in V} d_v = |E| = \alpha n^2$. מכיוון שהראינו שמתקיים $\mathrm{E}[|F_u|] \geq \alpha^2 n^2$, נובע מכך שקיימת בחירה ספציפית של u עבורה $|F_u| \geq \alpha^2 n^2$, כנדרש.

גלגלים מסתובבים

אנחנו נתייחס לקבוצת הנקודות V כאל קבוצת כל המספרים $\{0, \ldots, kr-1\}$ ב־$\mathbb{Z}_{kr}$, קבוצת השלמים מודולו kr. מעתה, כל החישובים שלנו בפתרון השאלה יהיו מודולו kr. בפרט, אם u, v, w שלוש נקודות ב־$\mathbb{Z}_{kr}$, ו־x, y, z שלוש נקודות אחרות, וכן מתקיים $x - u = y - v = z - w \pmod{kr}$, אז נובע מזה ששני המשולשים המתאימים חופפים (ע"י סיבוב).

עבור $V_1, \ldots, V_r$ נגריל עתה באופן יוניפורמי וב"ת אינדכסים $a_1, \ldots, a_r \in \mathbb{Z}_{kr}$, ולכל i נגדיר את הקבוצה $W_i = a_i + V_i = \{a_i + v : v \in V_i\}$, כאשר החיבור הוא מודולו kr. עבור $w \in \mathbb{Z}_{kr}$ כל שהוא, ההסתברות שמתקיים $w \in \bigcap_{i=1}^{r} W_i$ היא בדיוק $1/r^r$. זאת מכיוון שההסתברות ל־$w \in W_i$ היא $1/r$, ויש לנו חיתוך של r מאורעות ב"ת כאלה.

נגדיר משתני אינדיקטור X_w עבור המאורעות $w \in \bigcap_{i=1}^{r} W_i$, ונקבל $|\bigcap_{i=1}^{r} W_i| = \sum_{w \in \mathbb{Z}_{kr}} X_w$. מלינאריות התוחלת נקבל מכך שמתקיים $\mathrm{E}[|\bigcap_{i=1}^{r} W_i|] = k/r^{r-1}$. זה אומר שעבור $k > 2r^{r-1}$ התוחלת של $|\bigcap_{i=1}^{r} W_i|$ גדולה מ־2, ולכן קיימת בחירה ספציפית של $i_1, \ldots, i_r$ שעבורה גודל החיתוך הנ"ל גדול מ־2, ז"א שהוא מכיל לפחות שלוש נקודות שונות $\{x, y, z\}$. מתקבל מכך שלכל i הקבוצה V_i מכילה את המשולש $t_i = \{x - a_i, y - a_i, z - a_i\}$, ומשולשים אלו כולם חופפים.

קירבה לדרגה קבועה

ראשית נבצע ניתוח הסתברותי עבור d כללי, ואחר כך נבחר d שיתאים לנו. עבור שני צמתים $u, v \in V$, נסמן ב־N_{uv} את המאורע שיש ל־u לפחות d שכנים פרט ל־v, ונסמן ב־$E_{\{u,v\}}$ את המאורע ש־$\{u, v\}$ היא קשת (יש סימון מפורש של קבוצה כי אין מאורע נפרד "$E_{\{v,u\}}$", אבל לעומת זאת N_{uv} ו־N_{vu} הם כן מאורעות נפרדים). לבסוף נסמן את המאורע ה"רע" עבור זוג הצמתים ב־$A_{\{u,v\}} = E_{\{u,v\}} \wedge (N_{uv} \vee N_{vu})$.

ההסתברות של $E_{\{u,v\}}$ היא α/n בדיוק. התוחלת של מספר השכנים של u פרט ל־v (אם הוא שכן או לא) היא $(n-2) \cdot \alpha/n < \alpha$. על כן לפי אי שוויון מרקוב, ההסתברות של N_{uv} היא קטנה מ־α/d. לפי חסם איחוד מאורעות, ההסתברות של $N_{uv} \vee N_{vu}$ היא פחות מ־$2\alpha/d$. עתה נשים לב שהמאורע $E_{\{u,v\}}$ הוא ב"ת ב־$N_{uv} \vee N_{vu}$, בגלל שכל זוג שאינו $\{u, v\}$ נבחר להיות קשת באופן ב"ת לבחירה של האם $\{u, v\}$ קשת. על כן, ההסתברות עבור $A_{\{u,v\}}$, המאורע ש־$\{u, v\}$ היא קשת אשר לפחות אחד מצמתיה בעל דרגה גדולה מ־d (ז"א עם לפחות d שכנים מחוץ ל־$\{u, v\}$ עצמה), חסומה ע"י $2\alpha^2/dn$.

מכאן, לפי לינאריות התוחלת, תוחלת מספר הקשתות עם לפחות צומת אחד מדרגה גבוהה מ־d היא קטנה מ־$\alpha^2 n/d < (2\alpha^2/dn)\binom{n}{2}$. בחירה של $d = \alpha^2/\beta\gamma$ תיתן לנו (שוב לפי אי שוויון מרקוב) שבסיכוי יותר מ־$1-\gamma$ לא יהיו יותר מ־βn קשתות סמוכות לצמתים מדרגה גבוהה מ־d. במידה וזה קורה, אם אנחנו נסיר את כל הקשתות הנ"ל, אז לא יישארו קשתות סמוכות לצמתים מדרגה גבוהה מ־d, ולכן בפרט לא יהיו צמתים מדרגה כזו בגרף.

סיפוק חלקי של נוסחת CNF

ראשית, בלי הגבלת הכלליות ניתן להניח את ההנחות הבאות (שימו לב כי אנחנו כן מרשים לאותה פסוקית להופיע מספר פעמים בנוסחה):

- אין אף פסוקית מהצורה $\neg x_i$, שכן במקרה כזה נסמן $y_i \triangleq \neg x_i$ ונחליף את כל מופעי x_i ב $\neg y_i$. נזכר כי נתון שאם מופיעה הפסוקית $\neg x_i$, אז לא מופיעה הפסוקית x_i.
- בכל פסוקית המכילה ליטרל חיובי (כלומר לא בתוך שלילה) לא קיימים ליטרלים (חיוביים או שליליים) אחרים. ניתן להשיג זאת על ידי השמת סדר כלשהו על המשתנים, ולכל פסוקית בה יש מספר ליטרלים חיוביים בוחרים את המינימלי מבחינת הסדר ומסירים את כל הליטרלים האחרים מהפסוקית. ברור כי כל השמה שמספקת את הפסוקית החדשה מספקת גם את המקורית.
- כל פסוקית ללא ליטרלים חיוביים מורכבת משני ליטרלים שליליים בדיוק. שוב, אם יש יותר, ניתן לשמור רק את שני הליטרלים השליליים המינימליים מבחינת הסדר ולהסיר את כל הליטרלים האחרים מהפסוקית.

כעת ניתן להשתמש בלינאריות התוחלת. נבחר כל משתנה באופן בלתי תלוי להיות 1 בהסתברות $\alpha = \frac{1}{2}\left(\sqrt{5}-1\right)$ ולהיות 0 בהסתברות המשלימה $1-\alpha$. פסוקית מהצורה x_i תסתפק בהסתברות α. פסוקית מההצורה $\neg x_i \vee \neg x_j$ תסתפק בהסתברות

$$\begin{aligned}
1-\alpha^2 &= 1-\frac{1}{4}\left(\sqrt{5}-1\right)^2 \\
&= 1-\frac{1}{4}\left(5-2\sqrt{5}+1\right) \\
&= 1-\frac{3}{2}+\frac{1}{2}\sqrt{5} = \alpha
\end{aligned}$$

כך מלינאריות התוחלת תוחלת מספר הפסוקיות המסופקות היא $\frac{1}{2}\left(\sqrt{5}-1\right)$ מהפסוקיות, ולכן קיימת הצבה שמספקת לפחות כמספר הזה של פסוקיות.

זאב בודד

נגדיר את המשתנה המקרי X להיות מספר הצמתים חסרי השכנים בגרף שהוגרל, ונחשב את התוחלת שלו: ההסתברות של צומת מסויים להיות בודד היא

$$\begin{aligned}
(1-\frac{\ln(n)}{n})^{n-1} &\geq (e^{-\ln(n)/n-(\ln(n)/n)^2/2})^{n-1} \\
&= e^{-\ln(n)-(\ln(n))^2/2n}/e^{-\ln(n)/n-(\ln(n)/n)^2/2} \geq \frac{1}{n}e^{-(\ln(n))^2/2n} = (1-o(1))\frac{1}{n}
\end{aligned}$$

ולכן מתקיים לפי לינאריות התוחלת $\mathrm{E}[X] = 1-o(1)$. בשלב זה עדיין לא הוכחנו הרבה, כי X עלול לקבל ערכים גדולים עד n, אבל עתה נראה ש"הרבה מהתוחלת שלו" מתקבלת מערכים נמוכים. נשתמש בפיתוח $\mathrm{E}[X] = \sum_{i=1}^{\infty} i\Pr[X=i] = \sum_{k=1}^{\infty}\Pr[X \geq k]$.

אם מתקיים $X \geq k$, אז בפרט יש בגרף קבוצה U בת k צמתים ללא שכנים בין U ל־$V \setminus U$ (לוקחים את U להיות תת־קבוצה בת k איברים של קבוצת הצמתים חסרי השכנים). עבור $k < \ln(n)$ ו־$n \geq 16$ הסיכוי שזה קורה חסום

(דרך איחוד מאורעות) ע"י

$$\binom{n}{k}(1-\frac{\ln(n)}{n})^{k(n-k)} \leq \frac{n^k}{k!}e^{-k(n-k)\ln(n)/n} = \frac{n^k}{k!}\cdot n^{-k}\cdot e^{k^2\ln(n)/n} \leq \frac{1}{k!}e^{(\ln(n))^3/n} \leq \frac{2}{k!}$$

עבור $k \geq \ln(n)$, נחסום את ההסתברות של $X \geq k$ דרך ההסתברות למאורע המכיל אותו $X \geq \lceil \ln(n) \rceil$. עבור n גדול דיו (למשל 1024), מתקבל $\Pr[X \geq k] \leq \frac{1}{\lceil \ln(n)\rceil !}e^{\lceil \ln(n)\rceil^3/n} \leq \frac{1}{8n}$.

מאלו נובע שעבור n גדול דיו מתקיים

$$\frac{1}{2} \leq \mathrm{E}[X] = \sum_{k=1}^{\infty}\Pr[X \geq k] \leq \sum_{k=1}^{4}\Pr[X \geq k] + \sum_{k=5}^{\lceil \ln(n)\rceil -1}\frac{1}{k!} + (n-\lceil \ln(n)\rceil)\cdot\frac{1}{8n} \leq 4\Pr[X \geq 1] + \frac{1}{4}$$

ואז מהעברת אגפים נקבל $\Pr[X \geq 1] \geq \frac{1}{16} = \Omega(1)$, כנדרש.

הערה: יש עוד שיטות לבצע את החסמים, חלק עם שימוש ישיר (ללא פיתוח) בנוסחה $\mathrm{E}[X] = \sum_{i=1}^{\infty} i\Pr[X=i]$. חלקן יותר ארוכות מזו המוצגת כאן.

רב־קרב תרנגולים

נראה מה מתרחש בסוף סיבוב הניקורים הראשון: כל תרנגול ישאר בריא בהסתברות $\frac{1}{4}$, שזו ההסתברות לשני המאורעות הב"ת שגם התרנגול מימינו וגם התרנגול משמאלו לא ניקרו אותו. תרנגול יהפוך להיות בריא ובטוח אם גם לא ניקרו אותו, וגם בצד של התרנגול שהוא עצמו לא ניקר, השכן של השכן בחר כן לנקר את השכן. על כן יש הסתברות של $\frac{1}{8}$ להיות בריא ובטוח, והסתברות של $\frac{1}{8}$ להיות בריא ולא בטוח. מלינאריות התוחלת, תוחלת מספר התרנגולים הבריאים הבטוחים היא $n/8$, וזו גם תוחלת מספר התרנגולים הבריאים הלא־בטוחים.

נשים לב עתה שלאחר הסיבוב הראשון, בכל מקרה לכל תרנגול יהיה שכן מנוקר (השכן שהוא עצמו ניקר). על כן, כל התרנגולים הבריאים הלא־בטוחים יהיו בזוגות זה ליד זה, ללא שכנים בריאים אחרים. מספר הזוגות הוא בדיוק חצי ממספר התרנגולים המעורבים, ולכן התוחלת שלו היא $n/16$. עתה נרצה לחשב את תוחלת מספר התרנגולים ששורדים בסוף כל הסיבובים מבין הזוגות האלו. לכל זוג בודד מתקיים עתה תהליך זהה לזה של השאלה "דו־קרב תרנגולים" מהפרק על בניה וניתוח של מרחבי הסתברות, כי בכל סיבוב כל תרנגול יבחר באופן ב"ת האם לנקר את חברו לזוג, או את שכנו השני המנוקר כבר ממילא. נשתמש אם כן בתוצאת השאלה הנ"ל, ונציין שעבור כל זוג, הסיכוי שיישאר ממנו תרנגול בריא הוא $2/3$. על כן תוחלת מספר התרנגולים שהיו בריאים לא־בטוחים ונשארו בריאים עד סוף המשחק היא $2/3 \cdot n/16 = n/24$. על כן תוחלת מספר כל התרנגולים הבריאים (ע"י חיבור תוחלות שני המ"מ) היא $n/8 + n/24 = n/6$.

איזונים ובלמים

עבור הטענה שבהדרכה, $\mathrm{E}[|T_{i+1}|/|S_{i+1}|] = \mathrm{E}[|T_i|/|S_i|]$ (במקרה ש־$f|_{S_i}$ מקבל את שני הערכים האפשריים), דבר ראשון מראים שהיא מתקיימת לכל התניה אפשרית על זהות הקבוצה $S_i = A \subseteq S_0$ יחד עם ההתניה $|S_{i+1}| = k$ עבור $0 < k < |A|$ (אי־השוויון מבטיח שההתניה קורית בהסתברות חיובית).

תחת ההתניות האלו, הבחירה של S_{i+1} היא יוניפורמית מבין $\binom{|A|}{k}$ האפשרויות לת"ק מגודל k של $A = S_i$. לכל $j \in T_i$ (נשים לב שזהות $T_i = \{s \in A : f(s) = 1\}$ נקבעת כבר ע"י ההתניה שלנו על זהות S_i) נגדיר את המ"מ X_j להיות 1 אם j נבחר להיות ב־S_{i+1} ולהיות 0 אחרת. תחת הגדרות אלו מתקיים $|T_{i+1}| = \sum_{j\in T_i} X_j$ ולכן $\mathrm{E}\left[|T_{i+1}| \,\middle|\, S_i = A \wedge |S_{i+1}| = k\right] = \sum_{j\in T_i}\mathrm{E}[X_j] = |T_i|\cdot\frac{k}{|A|}$. לכן מתקיים $\mathrm{E}\left[|T_{i+1}|/|S_{i+1}| \,\middle|\, S_i = A \wedge |S_{i+1}| = k\right] = \frac{1}{k}\mathrm{E}\left[|T_{i+1}| \,\middle|\, S_i = A \wedge |S_{i+1}| = k\right] = |T_i|/|S_i|$ (השתמשנו בכך ש־$|S_{i+1}|$ שווה לערך הקבוע k תחת ההתניות).

מכיוון שהשוויון על התוחלת המותנה נכון עבור כל התניה על $|S_{i+1}|$, מתקבל לפי נוסחת התוחלת השלמה $\mathrm{E}\left[|T_{i+1}|/|S_{i+1}|\,\middle|\,S_i=A\right] = \sum_{k=1}^{|A|-1}\mathrm{E}\left[|T_{i+1}|/|S_{i+1}|\,\middle|\,S_i=A\wedge|S_{i+1}|=k\right]\Pr[|S_{i+1}|=k] = |T_i|/|S_i|$. מכך נובע $\mathrm{E}\left[|T_{i+1}|/|S_{i+1}|\right] = \sum_\alpha \mathrm{E}\left[|T_{i+1}|/|S_{i+1}|\,\middle|\,|T_i|/|S_i|=\alpha\right]\Pr[|T_i|/|S_i|=\alpha] = \mathrm{E}\left[|T_i|/|S_i|\right]$, כאשר הסכום באמצע הוא על כל ה־α שעבורם $\Pr[|T_i|/|S_i|=\alpha]>0$ (שזו תת־קבוצה סופית של קבוצת המספרים הרציונלים $\mathbb{Q}$).

לאחר שהוכחנו $\mathrm{E}\left[|T_{i+1}|/|S_{i+1}|\right] = \mathrm{E}\left[|T_i|/|S_i|\right]$, טיעון אינדוקטיבי מעל i נותן $\mathrm{E}\left[|T_i|/|S_i|\right] = \mathrm{E}\left[|T_0|/|S_0|\right]$, אבל זה עוד לא מסיים את ההוכחה, כי ה־i שבו התהליך נעצר אינו קבוע מראש.

על מנת להמשיך, "נרחיב" את התהליך כך שתמיד יסתיים ב־$i=|S_0|$. לכל $i<|S_0|$, אם הפונקציה $f|_{S_i}$ אינה קבועה אז נעשה בדיוק את מה שעשינו לפי התהליך המקורי. אם הפונקציה היא כבר קבועה, אז במקום לעצור מיידית, נקבע בפשטות $S_{i+1}=S_i$ (בהסתברות 1). נשים לב ש־$f|_{S_{|S_0|}}$ היא בהכרח פונקציה קבועה, כי בכל שלב i שבו הפונקציה אינה קבועה בהכרח יתקיים $|S_{i+1}|<|S_i|$, ולא יתכן שהגודל של S_i יתאפס. בשלב האחרון נפלוט את הערך הקבוע של f (שזה אותו ערך שבאלגוריתם המקורי היינו פולטים קודם לכן).

הדבר הבא לשים לב הוא שגם בתהליך החדש מתקיים $\mathrm{E}\left[|T_{i+1}|/|S_{i+1}|\right] = \mathrm{E}\left[|T_i|/|S_i|\right]$ לכל $i<|S_0|$. לשם כך מפצלים למקרים. את המקרה שבו $f|_{S_i}$ אינה קבועה כבר ניתחנו קודם, והמקרה שבו זאת פונקציה קבועה (ואז $S_{i+1}=S_i$) הוא מיידי (כי אז מתקיים $|T_{i+1}|/|S_{i+1}|=|T_i|/|S_i|$ גם בלי לקיחת תוחלת). בסוף מתקיים $|T_{|S_0|}|/|S_{|S_0|}|\in\{0,1\}$ (כי f קבועה על $S_{|S_0|}$, וזה גם יהיה הפלט הסופי של האלגוריתם), ומאינדוקציה מתקיים $\mathrm{E}\left[|T_{|S_0|}|/|S_{|S_0|}|\right] = \mathrm{E}\left[|T_0|/|S_0|\right]$, ומכאן שהסיכוי שהערך הנ"ל יהיה שווה ל־1 הוא בדיוק $|T_0|/|S_0|$.

לבלוע את החוכמה

נגדיר את מרחב ההסתברות τ על זוגות של ת"ק $Q,Q'\subseteq\{1,\ldots,n\}$ באופן הבא:

- ראשית נגריל את Q לפי ν, ז"א לכל $i\in\{1,\ldots,n\}$ נבחר אותו להיות איבר ב־Q בהסתברות p, באופן ב"ת בבחירות לאינדכסים אחרים ב־$\{1,\ldots,n\}$.
- אם מתקיים $|Q|\ge q$, אז נקבע את $Q'=Q$.
- אם מתקיים $|Q|<q$, אז נגריל את $R\subseteq\{1,\ldots,n\}\setminus Q$ באופן יוניפורמי מכל תתי־הקבוצה האפשריים בגודל $q-|Q|$, ואז נקבע $Q'=Q\cup R$.

נסמן ב־A_1 את המאורע "קיימת $F\in\mathcal{F}$ שעבורה מתקיים $F\subseteq Q$", וב־A_2 את המאורע "קיימת $F\in\mathcal{F}$ שעבורה מתקיים $F\subseteq Q'$". ראשית נשים לב שמתקיים $\Pr_\tau[A_2]\ge\Pr_\tau[A_1]$, מכיוון שתמיד מתקיים $Q\subseteq Q'$. כמו כן מתקיים $\Pr_\tau[A_1]=\Pr_\nu[A]$ כי Q מתפלג באופן זהה תחת τ ו־ν.

עתה נסמן ב־B את המאורע "$|Q|>q$". נשים לב שמתקיים $\Pr_\tau[B]<\frac{pn}{q}$ לפי אי שוויון מרקוב (מתקיים $\mathrm{E}_\tau[|Q|]=pn$ לפי לינאריות התוחלת, כי אם נגדיר את X_i להיות משתנה האינדיקטור עבור "$i\in Q$", אז נקבל $|Q|=\sum_{i=1}^n X_i$). כמו כן, לכל $1\le t\le q$, ההתפלגות של Q' תחת התנאי "$|Q|=t$" זהה להתפלגות ν, מכיוון שאז, לכל $T\subseteq\{1,\ldots,n\}$ מגודל q, מחישוב ישיר מתקבל $\Pr_\tau[Q'=T\,|\,|Q|=t]=\binom{q}{t}/\binom{n}{t}\binom{n-t}{q-t}=1/\binom{n}{q}$. לכן ההתפלגות של Q' תחת τ בהינתן המאורע $\neg B$ (שהוא איחוד המאורעות "$|Q|=t$" עבור $1\le t\le q$) זהה להתפלגות של Q תחת μ.

מאלו מתקיים $\Pr_\mu[A]=\Pr_\tau[A_2|\neg B]\ge\Pr_\tau[A_2\wedge\neg B]\ge\Pr_\tau[A_2]-\Pr_\tau[B]>\Pr_\tau[A_2]-\frac{pn}{q}$, ולסיכום נקבל $\Pr_\mu[A]>\Pr_\tau[A_2]-\frac{pn}{q}\ge\Pr_\tau[A_1]-\frac{pn}{q}=\Pr_\nu[A]-\frac{pn}{q}$ כנדרש.

כל אחד את עצמו

נסמן ב־V_i את קבוצת הצמתים בשלב ה־i שיש להם שכן שצבוע באותו צבע כמותם. בפרט, אם הגרף הוא חסר צמתים מבודדים אז $V_0=V$ (כי אז כל הצמתים צבועים באותו צבע), אבל בכל מקרה $|V_0|\le n$. שימו לב ש־$|V_i|$

(לכל $i \geq 0$) הוא משתנה מקרי, כי הוא תלוי בזהות הצמתים החברים ב־V_i, ואלו תלויים בתהליך המקרי של האלגוריתם.

נוכיח עתה באינדוקציה שמתקיים $\mathrm{E}[|V_k|] \leq (\frac{2}{3})^k n$. הבסיס הוא $k=0$. עבור המעבר מהשלב ה־$k-1$ לשלב ה־k, נחסום מלמעלה את ההסתברות של צומת להיות עם שכן מאותו צבע. נסמן $r=|V_{k-1}|$ (לפי הנחת האינדוקציה מתקיים $\mathrm{E}[r] \leq (\frac{2}{3})^{k-1} n$), ונסמן $V_{k-1}=\{u_1,\ldots,u_r\}$. הסדר שבו אנחנו מסמנים את הצמתים יכול להיות שרירותי, זה לא משנה אפילו אם נשתמש בסדר אחר לכל איטרציה של האלגוריתם.

עבור הניתוח נניח שאנחנו בוחרים את הצבעים החדשים של הצמתים לפי הסדר השרירותי הנ"ל, החל מ־u_1 וכלה ב־u_r. נחסום את תוחלת הגודל של W, קבוצת הצמתים שהפכו להיות עם שכן מאותו צבע בשלב כל שהוא של הצביעה של V_{k-1}, שזה בוודאי חוסם את מספר הצמתים ב־$V_k \subseteq W$ (יכול להיות שחלק מהצמתים של W "ניצלו" כשצבענו מחדש צמתים יותר מאוחרים, אבל בכל מקרה יש הכלה).

כאשר אנחנו צובעים את u_j מחדש, נתון שיש לו לא יותר מ־d שכנים. מכיוון שהצבע של u_j נבחר באופן יוניפורמי (וב"ת בבחירות קודמות) מקבוצה בת $3d$ צבעים, לפי לינאריות התוחלת, תוחלת מספר השכנים שלו שיהיו כתוצאה מזה מאותו צבע חסומה ע"י $\frac{1}{3}$ (כל שכן יהיה בצבע זהה בהסתברות $\frac{1}{3d}$). צריך אבל לזכור שכאשר יש שכן מאותו צבע אז גם u_j עצמו הוא בעל שכן מצבע זהה, ולכן תוחלת מספר הצמתים הכולל שיהיה להם שכן מאותו צבע בגלל הצביעה של u_j חסומה ע"י $\frac{2}{3}$. שוב מלינאריות התוחלת, מקבלים שתוחלת גודל הקבוצה W חסום כולו ע"י $\frac{2}{3}r$. כאשר אנחנו לא מתנים על r, נקבל חסם עבור התוחלת הלא מותנה $\mathrm{E}[|V_k|] \leq \mathrm{E}[|W|] \leq \frac{2}{3}\mathrm{E}[|V_{k-1}|]$, ומכאן אנחנו משלימים את צעד האינדוקציה עבור $\mathrm{E}[|V_k|] \leq (\frac{2}{3})^k n$.

נסמן ב־T את המ"מ של הזמן שבו כל הצמתים היו צבועים בצבעים שונים. לפי אי־שוויון מרקוב מתקבל $\Pr[T>k] = \Pr[|V_k| \geq 1] \leq (\frac{2}{3})^k n$. כמו כן, עבור $k \leq 2\log n$, בוודאי מתקיים $\Pr[T>k] \leq 1$. מכאן, ע"י שימוש בטריק $\sum_{k=0}^{\infty} k\alpha_k = \sum_{k=0}^{\infty}(\sum_{j=k+1}^{\infty}\alpha_j)$ כאשר אצלנו $\alpha_k = \Pr[T=k]$, אפשר להשיג את החסם $\mathrm{E}[T] = \sum_{k=0}^{\infty}\Pr[T>k] \leq 2\log n + \sum_{k \geq 2\log n}(\frac{2}{3})^k n \leq 2\log n + \sum_{j=0}^{\infty}(\frac{2}{3})^j = O(\log n)$.

לתפוס את המרובה

נחשוב על סדרה לא סופית של מ"מ $X_1, X_2, \ldots$ שכולם ב"ת ומתפלגים לפי μ. לשם הנוחות, לכל $i \in \mathbb{N}$ נגדיר את הקבוצה (המוגרלת מקרית) A_i להיות קבוצת כל הערכים המופיעים ב־$\{X_1,\ldots,X_i\}$, ובאופן תואם לכך נגדיר גם את $A_0=\emptyset$. בנוסף לכך לכל $j \in \mathbb{N}$ נגדיר את K_j להיות המ"מ שמקבל את ה־i המינימלי שעבורו $|A_i|=j$, ובהתאם נגדיר גם $K_0=0$. יכול להיות מצב שיתקיים $K_j=\infty$: זה תמיד קורה אם ל־μ יש תומך סופי שגודלו קטן מ־j, ולא קשה לראות שזה קורה בהסתברות 0 אחרת. עתה נגדיר את המשתנים שננתח, $Z_1,\ldots,Z_n$.

לכל $1 \leq i \leq n$, Z_i יוגדר להיות שווה ל־0 אם קיים $j \leq K_{i-1}$ שעבורו $\mu(A_j) \geq 1-\epsilon$ (הסיבה היחידה לא להתייחס ישר ל־$A_{K_{i-1}}$ היא מקרה הקצה שמתקיים $K_{i-1}=\infty$), ואחרת הוא יוגדר להיות שווה ל־$K_i - K_{i-1}$, ז"א "מספר הנסיונות להגדיל את $|A_{K_{i-1}}|$ מ־$i-1$ ל־i עד שמצליחים". עתה נשים לב שהסכום $Z=\sum_{i=1}^{n} Z_i$ הוא המ"מ שערכו שווה ל־j הקטן ביותר שעבורו A_j מקיים את תנאי השאלה "$\mu(A_j) \geq 1-\epsilon$ או $|A_j| \geq n$". הסיבה לכך היא שהסכום הנ"ל שווה ל־K_n אם לא היה $j<K_n$ שעבורו $\mu(A_j) \geq 1-\epsilon$, ואחרת הוא שווה ל־K_i שעבורו $j=K_i$ הוא האינדכס שבו האירוע "$\mu(A_j) \geq 1-\epsilon$" קרה לראשונה. לכן, אם מתקיים $\mathrm{E}[Z] \leq \alpha$, אז אפשר לפי אי־שוויון מרקוב לבחור $k=\lceil 2\alpha \rceil$ ולקבל חסם הסתברות של לפחות $\frac{1}{2}$ ש־A_k יקיים את תנאי השאלה.

נוכיח שלכל i מתקיים $\mathrm{E}[Z_i] \leq 1/\epsilon$: נראה שלכל קבוצה סופית $B \subset \mathbb{N}$, מתקיים $\mathrm{E}[Z_i|A_{K_{i-1}}=B] \leq 1/\epsilon$ (אלא אם כן מתקיים $\Pr[A_{K_{i-1}}=B]=0$). בפרט הדבר יוכיח שכל ה־Z_i יהיו סופיים בהסתברות 1 (שימו לב שהגדרת Z_i לא מבטיחה את זה אוטומטית). אם $\mu(B) \geq 1-\epsilon$, אז מההגדרה מתקיים $\mathrm{E}[Z_i|A_{K_{i-1}}=B]=0$. לעומת זאת, אם $\mu(B) = 1-\eta < 1-\epsilon$, אז ההתפלגות של Z_i זהה למספר ההטלות של מטבע לא מאוזן עם הסתברות η לתוצאה "1", עד שתוצאה זו מתקבלת (ההתפלגות הגיאומטרית עם פרמטר η), ובפרט מתקיים $\mathrm{E}[Z_i|A_{K_{i-1}}=B] = 1/\eta \leq 1/\epsilon$. מכיוון שהחסם נכון לכל התניה אפשרית על $A_{K_{i-1}}$, מתקיים גם ללא התניה

$\mathrm{E}[Z_i] \le 1/\epsilon$. על כן מלינאריות התוחלת מתקיים $\mathrm{E}[Z] \le n/\epsilon$, ולכן בחירה של $k = 2n/\epsilon$ תקיים את הנדרש עבור פתרון השאלה.

הערה: הניתוח של $\mathrm{E}[Z]$ מזכיר את השאלה "הגעה מהוססת".

לא כל הדרכים מובילות

נניח שלכל v מתקיים $\mathrm{E}[T_v] < n/4$, ונראה סתירה. לפי אי שוויון מרקוב מתקיים אז $\Pr[T_v \ge n/2] < \frac{1}{2}$, ז"א שבסיכוי גדול מ־$\frac{1}{2}$ הגענו לצומת v בפחות מ־$n/2$ צעדים. נסמן ב־I_v את משתנה האינדיקטור עבור המאורע הזה, וב־Y את המ"מ המקבל את מספר הצמתים השונים שביקרנו בהם בפחות מ־$n/2$ צעדים. מתקיים $Y = \sum_{v \in V} I_v$ כאשר V מסמן את קבוצת הצמתים של הגרף, ולכן $\mathrm{E}[Y] = \sum_{v \in V} \mathrm{E}[I_v] > n \cdot \frac{1}{2}$. זוהי סתירה לעובדה שערכו של Y לעולם אינו יכול לעלות על $n/2$ (גם אם כוללים את s), כי בכל צעד אנחנו לא מבקרים ביותר מצומת חדש אחד.

דה־רנדומיזציה

פסוקיות בסדר

בשאלה זו נדבר על מערכת של m פסוקיות, כמו מערכות 3CNF שנידונו בשיעור, אבל כאן, במקום למצוא הצבה למשתנים בוליאנים, צריך למצוא פרמוטציה (פונקציה חח"ע ועל) $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, והפסוקיות מתייחסות לסדר שהפרמוטציה קובעת. ספציפית, כל פסוקית היא מהצורה "$\neg(\sigma(i) < \sigma(j) < \sigma(k))$" עבור $1 \le i, j, k \le n$ שונים זה מזה. הכוונה היא שהפסוקית הזו מסתפקת בכל מקרה **למעט** המקרה שבו הפרמוטציה σ גם נתנה ל־i ערך קטן מזה שנתנה ל־j, וגם נתנה לשני אלו ערך קטן מזה שנתנה ל־k. הראו שקיימת פרמוטציה שמספקת בו זמנית לפחות $\frac{5}{6}m$ מתוך m הפסוקיות, וכתבו אלגוריתם דטרמיניסטי מפורט (עם זמן ריצה פולינומי ב־m ו־n) שמוצא פרמוטציה כזו.

בלתי תלויים בשלשות

הראו עבור $k \ge 1$ שאפשר לבנות 2^k משתנים מקריים, אשר מקבלים כ"א ערך יוניפורמי מתוך $\{0, 1\}$ וכך שכל שלושה מהם הם בלתי תלויים, כך שגודל מרחב ההסתברות כולו הוא 2^{k+1} בלבד.

מרחב מדגם מוגבל מוטה

אנחנו מעוניינים במרחב הסתברות שעבורו מוגדרים משתנים מקריים $X_1, \ldots, X_n$, כך שלכל i מתקיימים השוויונים $\Pr[X_i = 0] = \frac{2}{3}$ ו־$\Pr[X_i = 1] = \frac{1}{3}$, וכן המשתנים הנ"ל הם ב"ת בזוגות (לכל $i < j$ מתקיים ש־X_i ב"ת ב־X_j). הראו שיש מרחב כזה שמספר האיברים הכולל בו הוא פולינומי ב־n.

פתרונות לתרגילים על דה־רנדומיזציה

פסוקיות בסדר

ההוכחה שקיימת הצבה שמספקת לפחות $\frac{5}{6}m$ פסוקיות מתבצעת באמצעות לינאריות התוחלת. מגרילים את σ באופן מקרי ויוניפורמי (מתוך $n!$ האפשרויות לפרמוטציה הזו), מסמנים ב־X_r עבור $1 \le r \le m$ את משתנה האינדיקטור עבור המאורע שהפסוקית ה־r מסתפקת, ומסמנים ב־$X = \sum_{r=1}^{m} X_r$ את המ"מ עבור מספר הפסוקיות המסתפקות.

מכיוון שסדרת הערכים $\sigma(i), \sigma(j), \sigma(k)$ מקבלת את כל אחד מ־$3! = 6$ הסדרים האפשריים בהסתברות זהה (צמצום של סדר מקרי שנבחר יוניפורמית לחלק מהאיברים הוא בעצמו סדר מקרי יוניפורמי), ויש רק מקרה אחד

שבו הפסוקית לא מסתפקת, מקבלים $\mathrm{E}[X_r] = \frac{5}{6}$ לכל $1 \leq r \leq m$, ולכן $\mathrm{E}[X] = \frac{5}{6}m$. מכאן שבהסתברות גדולה מ־0 מתקיים $X \geq \frac{5}{6}m$ כנדרש.

על מנת למצוא פרמוטציה כזו באופן דטרמיניסטי, נשתמש בשיטת התוחלות המותנות. הנוסחאות לתוחלות המותנות יהיו פשוטות יותר אם נבחר בשלבים את הפרמוטציה ההפוכה σ^{-1} במקום את σ. בשלב הראשון נבחר את $\sigma^{-1}(1)$, ז"א את i_1 עבורו מתקיים $\sigma(i_1) = 1$, בשלב השני את $\sigma^{-1}(2)$, וכו'. בשלב ה־s נבחר את $i_s = \sigma^{-1}(s)$, ונבחר את זה שיתן מקסימום לתוחלת המותנת המתאימה $\mathrm{E}[X|\sigma(i_1) = 1, \ldots, \sigma(i_s) = s]$. את החישוב הזה עושים לפי הסכום $\sum_{r=1}^{m} \mathrm{E}[X_r|\sigma(i_1) = 1, \ldots, \sigma(i_s) = s]$.

חישוב הסתברות מותנה בודדת $\Pr[\neg(\sigma(i) < \sigma(j) < \sigma(k))|\sigma(i_1) = 1, \ldots, \sigma(i_s) = s]$ נעשה באופן הבא:

- אם אף אחד מהמספרים i, j, k אינו איבר בסדרה $i_1, \ldots, i_s$, ההסתברות לסיפוק היא עדיין $\frac{5}{6}$.
- אם רק i איבר בסדרה, אז מכיוון ש־$\sigma(i)$ בטוח יהיה יותר קטן מ־$\sigma(j)$ ו־$\sigma(k)$ שעוד לא נבחרו, ההסתברות לסיפוק היא $\frac{1}{2}$.
- אם i אינו איבר ב־$i_1, \ldots, i_s$ אבל לפחות אחד מ־j ו־k כן איבר שם, אז ההסתברות לסיפוק היא 1, כי בטוח ש־$\sigma(i)$ לא יהיה קטן משני הערכים האחרים. באופן דומה, אם j אינו איבר בסדרה אבל k כן איבר בה, אז ההסתברות לסיפוק היא 1.
- אם i נמצא בסדרה לפני j (ז"א שבטוח $\sigma(i) < \sigma(j)$) ו־k אינו בסדרה (ואז הערך $\sigma(k)$ יהיה גדול מהם), אז ההסתברות לסיפוק היא 0.
- אם i נמצא בסדרה אחרי j (ואז $\sigma(i) > \sigma(j)$) אז ההסתברות היא 1. מצב דומה קורה אם k נמצא בסדרה כאשר לפחות אחד מ־i ו־j נמצאים בסדרה אחריו.
- המקרה היחידי שנותר הוא כאשר i נמצא בסדרה לפני j בעוד ש־k נמצא בסדרה אחרי שניהם, ואז הסתברות הסיפוק היא 0.

הערה: היה אפשר לגשת בשיטה יותר "ברוטלית" לחישוב התוחלות המותנות. למשל, אם בוחרים את σ במקום את σ^{-1} איבר־איבר, אז במקום לתת נוסחה "סגורה" עבור $\mathrm{E}[X_r|\sigma(1) = i_1, \ldots, \sigma(s) = i_s]$ (שהיא טיפה יותר מסובכת מזו שיוצאת מבחירת σ^{-1} שנכתבה למעלה), אפשר פשוט "לרוץ" על כל הערכים האפשריים עבור המשתנים ה"פנויים" בפסוקית ה־r (יש לא יותר משלושה משתנים כאלו, ואם שלושתם פנויים אז התוחלת המותנית היא $\frac{5}{6}$ ללא צורך בחישוב נוסף), ולספור בכמה מהאפשרויות הפסוקית מסתפקת. זה עדיין עונה על דרישות השאלה (כי דרשנו זמן ריצה פולינומי אבל לא הגבלנו את דרגת הפולינום), אבל בזבזני למדי – זמן הריצה יהיה $O(mn^4)$, בעוד שהאלגוריתם למעלה הוא בעל זמן ריצה $O(mn^2)$ באופן שנכתב.

בלתי תלויים בשלשות

נסמן ב־$Y_1, \ldots, Y_{k+1}$ סדרה של משתנים מקריים ב"ת לחלוטין שמקבלים ערכים באופן יוניפורמי מ־$\{0, 1\}$. אלו יהוו את מרחב ההסתברות שלנו. עתה נסמן ב־$V \subset \{0, 1\}^{k+1}$ את קבוצת כל הווקטורים הבינארים מאורך $k+1$ שלהם מספר אי־זוגי של ערכי 1. לא קשה לראות שמתקיים $|V| = 2^k$. עתה, לכל $v \in V$ נגדיר את המשתנה המקרי $X_v = \bigoplus_{i=1}^{k+1} v_i Y_i$, כאשר נסמן $v = (v_1, \ldots, v_{k+1})$. נראה עתה שאלו בלתי־תלויים בשלשות.

ההוכחות שאלו משתנים מקריים יוניפורמים וב"ת בזוגות כבר נעשו למעשה בפרק התרגול על מרחבי מדגם מוגבלים. הדבר העיקרי לשים לב עתה הוא שלכל $u, v \in V$, הסכום $u \oplus v$ שלהם (מודולו 2) מכיל מספר זוגי של ערכי 1. על כן לא יהיה ווקטור ב־V ששווה ל־$u \oplus v$, ומכאן שלכל $w \in V$ השונה מ־u ו־v, ערך X_w יוגרל באופן בלתי־תלוי מהערך של $(X_u \oplus X_v) = \bigoplus_{i=1}^{k+1} (u_i \oplus v_i) Y_i$. מאי תלות זו יחד עם אי התלות של X_u ו־X_v זה בזה (כאשר $u, v, w \in V$ כולם שונים זה מזה) נובעת אי התלות של משתנים אלו כשלישיה.

מרחב מדגם מוגבל מוטה

ראשית נבנה מרחב הסתברות עם משתנים מקריים $Z_1,\ldots,Z_n$, כך שכל Z_i מתפלג יוניפורמית מעל $\{0,1,2\}$, וכן כל המשתנים הנ"ל הם ב"ת בזוגות. מאלו אפשר לבנות את $X_1,\ldots,X_n$ ע"י כך שנקבע $X_i=1$ אם $Z_i=0$ ואחרת $X_i=0$. עבור בחירת $Z_1,\ldots,Z_n$, ניקח $k=\lfloor\log_2 n\rfloor+1$, נגריל את $Y_1,\ldots Y_k$ להיות משתנים מקריים יוניפורמים ב־$\{0,1,2\}$ וב"ת לחלוטין, ואז לכל קבוצה $\emptyset\neq A\subseteq\{1,\ldots,k\}$ נגדיר את $Z_A=\bigoplus_{a\in A}Y_a$, כאשר כאן $\bigoplus$ מסמן סכום מודולו 3. לבסוף, נגדיר $Z_i=Z_{A_i}$ כאשר $A_1,\ldots A_n$ הן קבוצות לא־ריקות שונות זו מזו.

גודל מרחב ההסתברות הוא $3^k=O(n^{\log_2 3})$, וזה פולינומי ב־n. ההוכחה שכל מ"מ Z_A מתפלג יוניפורמית ב־$\{0,1,2\}$ מאוד דומה לזו שנעשתה בתרגול עבור מרחבי דגימה מוגבלים, ולא נציג אותה מחדש כאן.

באשר לאי־תלות, נראה למשל שמתקיים $\Pr[Z_A=0|Z_B=0]=\frac{1}{3}$ לכל $A\neq B$. לצורך זה נניח שקיים איבר $b\in B\setminus A$ (המקרה שבו קיים איבר ב־$A\setminus B$ הוא בעל הוכחה זהה). נניח לשם פישוט הביטויים שנכתוב שמתקיים גם $b=1$, כמובן שההוכחה תהיה אותו דבר ל־b אחרים. נשים עתה לב שלכל $\beta_2,\ldots,\beta_k$ מתקיים:

$$\Pr[Z_A=0|Y_2=\beta_2,\ldots,Y_k=\beta_k]=\Pr\Big[Y_1=3-\bigoplus_{a\in A\setminus\{1\}}\beta_a\Big|Y_2=\beta_2,\ldots,Y_k=\beta_k\Big]=\frac{1}{3}$$

מכאן אפשר לסיים לפי נוסחת ההסתברות השלמה (תוך שימוש בכך שערך Z_B נקבע ע"י ערכי $Y_2,\ldots,Y_k$).

$$\begin{aligned}\Pr[Z_A=0|Z_B=0] &= \sum_{\beta_2,\ldots,\beta_k}\Pr[Z_A=0|Y_2=\beta_2,\ldots,Y_k=\beta_k]\Pr[Y_2=\beta_2,\ldots,Y_k=\beta_k|Z_B=0]\\ &= \sum_{\beta_2,\ldots,\beta_k}\frac{1}{3}\Pr[Y_2=\beta_2,\ldots,Y_k=\beta_k|Z_B=0]=\frac{1}{3}\end{aligned}$$

הגרלה עם תיקונים

קבוצות ב"ת בהיפרגרפים

עבור היפרגרף 3־יוניפורמי (ז"א מבנה עם קבוצת צמתים V וקבוצת "קשתות" E שבה כל קשת היא תת קבוצה של V בת שלושה צמתים בדיוק) בעל n צמתים ו־m קשתות, כאשר $m\geq\frac{1}{3}n$, הראו כי קיימת קבוצת צמתים בלתי תלויה (ז"א קבוצה $V'\subseteq V$ שאינה מכילה אף קשת מ־E) שגודלה לפחות $\frac{2n^{3/2}}{3\sqrt{3m}}$.

מספרי רמזי לא סימטרים

נסמן ב־$R(4,k)$ את מספר הצמתים המכסימלי שעבורו אפשר לבנות גרף שאינו מכיל קליק עם 4 צמתים או קבוצה ב"ת בת k צמתים. הראו כי $R(4,k)\geq\Omega((k/\log k)^2)$.

נזכיר אי שוויון שיכול לעזור כאן ובשאלות אחרות על גרפים: $(\frac{n}{k})^k\leq\binom{n}{k}<(\frac{en}{k})^k$, עבור $1\leq k\leq n$.

שיפודים

הראו שלכל היפרגרף k־יוניפורמי מעל קבוצת הצמתים $V=\{1,\ldots,n\}$ (כזכור כל קשת של היפרגרף כזה היא תת־קבוצה של V מגודל k בדיוק) שיש לו m קשתות ניתן למצוא קבוצה $U\subset V$ בת לכל היותר $(n\ln k+m)/k$ צמתים (שימו לב ללוגריתם בבסיס טבעי בביטוי), כך שכל קשת של ההיפרגרף תכיל לפחות צומת אחד מ־U. עליכם להסביר גם כיצד למצוא קבוצה כזו באמצעות אלגוריתם **דטרמיניסטי** עם זמן ריצה פולינומי ב־n ו־m.

דרגה, צביעה, מותן

הראו לכל k קבוע, ולכל n גדול מספיק (ביחס ל־k), שאפשר למצוא גרף עם n צמתים, ודרגה חסומה ע"י קבוע d (תלוי ב־k), כך שאין לו k־צביעה וגם אין בו מעגלים מגודל קטן מ־$C \log n$, עבור קבוע $C > 0$ מתאים (גם תלוי ב־k). אפשר לעשות את זה תוך שימוש בניתוח של השאלה "קירבה לדרגה קבועה" מהפרק על לינאריות התוחלת.

רוקדים על שתי החתונות

הראו, עבור d ו־n גדולים מספיק, שלכל גרף (פשוט ולא מכוון) G בעל קבוצת צמתים V מגודל n ובעל דרגה מינימלית לפחות d, קיימת קבוצה $A \subset V$ מגודל $O(n \log(d)/d)$, כך שלכל צומת $v \in V \setminus A$ קיים לפחות שכן אחד ב־A ולפחות שכן אחד ב־$V \setminus A$.

הבהרה: עליכם להוכיח, עבור קבוע גלובלי C מתאים, את הקיום של קבוצה A שגודלה חסום ע"י $Cn \log(d)/d$ ומקיימת את תנאי השאלה בנוגע לצמתים ב־$V \setminus A$.

פתרונות לתרגילים על הגרלה עם תיקונים

קבוצות ב"ת בהיפרגרפים

נסתכל על הפרוצדורה הבאה: ראשית נגריל קבוצת צמתים U ע"י כך שכל צומת ב־V יבחר באופן ב"ת בהסתברות α (את ערכו של α נבחר אח"כ). עתה נקבל ממנה קבוצת צמתים ב"ת W ע"י כך שמכל קשת של ההיפרגרף המוכל ב־U נחסר את אחד מצמתיה. אם נסמן ב־X את מספר הצמתים ב־U וב־Y את מספר הקשתות המוכלות ב־U, הרי שגודל W הוא לפחות $X - Y$ (יתכן שהוא גדול יותר). נחשב אם כן את תוחלת הפרש זה: $\mathrm{E}[X - Y] = \mathrm{E}[X] - \mathrm{E}[Y] = \alpha n - \alpha^3 m$. עתה נבחר את ה־$\alpha$ שלנו: ע"י גזירה לפי α וחיפוש נקודה המאפסת את הנגזרת נקבל $\alpha = \sqrt{\frac{n}{3m}}$ (כאן חשוב שיתקיים $m \geq \frac{1}{3}n$ כדי שנקבל $\alpha \leq 1$). ע"י הצבה נקבל עבור ערך זה $\mathrm{E}[X - Y] = \frac{2n^{3/2}}{3\sqrt{3m}}$, ומכאן שקיימת בחירה ספציפית של U שעבורה הפרש מספר הצמתים ומספר המשולשים אכן אינו יורד מביטוי זה. הקבוצה W שנקבל מ־U תקיים אם כן את המבוקש.

מספרי רמזי לא סימטרים

נסתכל על הגרף G בעל n הצמתים שבו כל זוג נבחר להיות קשת באופן ב"ת בהסתברות $n^{-1/2}$. תוחלת מספר העותקים של K_4 (הגרף השלם בעל 4 צמתים) בגרף זה היא $\binom{n}{4}(n^{-1/2})^6 < \frac{n}{12}$, ולכן בהסתברות לפחות $\frac{5}{6}$ קיימים ב־G לא יותר מ־$\frac{n}{2}$ עותקים שונים של K_4. בנוסף, אם $n = \lfloor \frac{1}{16}(\frac{k}{\ln k})^2 \rfloor$ (כאשר הלוגריתם כאן הוא בבסיס טבעי), אז הסיכוי שיש ב־G קבוצת צמתים ב"ת כל שהיא בגודל k חסום (עבור k גדול דיו) ע"י

$$\binom{n}{k}(1 - n^{-1/2})^{\binom{k}{2}} < (\frac{en}{k})^k e^{-n^{-1/2}\binom{k}{2}} < (ek)^k e^{-2(k-1)\ln k} = e^{k(1+\ln k)-2(k-1)\ln k} = o(1)$$

ולכן עבור כל k גדול דיו קיים גרף G בעל n צמתים שבו אין קבוצה ב"ת מגודל k וכן אין יותר מ־$\frac{n}{2}$ עותקים של K_4.

עתה נבחר את G' להיות הגרף המתקבל מ־G ע"י כך שלכל עותק של K_4 נסיר את אחד מצמתיו מ־G. ב־G' אין לא עותקים של K_4 ולא קבוצות ב"ת מגודל k, ומספר צמתיו הוא לפחות $\frac{1}{2}n = \Omega((k/\ln k)^2)$.

שיפודים

ראשית נראה שיקול הסתברותי (שמשתמש בהגרלה עם תיקונים) שמוכיח שיש קבוצה U המקיימת את הנדרש, ואחר כך נראה איך ניתן לבצע לו דה־רנדומיזציה בשיטת התוחלות המותנות. אנחנו נגריל קבוצה U_1 באופן הבא: כל צומת $v \in V$ (באופן ב"ת בצמתים האחרים) יוכנס ל־U_1 בהסתברות $\frac{\ln(k)}{k}$ בדיוק. נגדיר מ"מ Y_1 שיהיה שווה

לגודל של U_1, ונגדיר מ"מ Y_2 שיהיה שווה למספר הקשתות של ההיפרגרף שאינן מכילות צומת מ־U_1. על מנת להגדיר את U, נוסיף ל־U_1 צומת שרירותי מכל קשת שאינה מכילה צומת מ־U_1, ואז נקבל קבוצה כך שכל קשת מכילה צומת ממנה ומקיימת $|U| \leq Y_1 + Y_2$ (זה לא בהכרח שוויון כי יכול להיות שבשלב האחרון היו צמתים שהוספנו "מספר פעמים" עבור מספר קשתות של ההיפרגרף).

עתה נחסום את התוחלות של המ"מ, לפי לינאריות התוחלת. מתקיים $\mathrm{E}[Y_1] = n\frac{\ln(k)}{k}$ (לפי כך שלכל צומת $v \in V$ מגדירים משתנה אינדיקטור עבור המאורע שהוא נכנס ל־U_1, והתוחלת של מ"מ זה היא $\frac{\ln(k)}{k}$). כמו כן, לכל קשת של ההיפרגרף, הסיכוי שהיא אינה מכילה צמתים מ־U_1 הוא $(1 - \frac{\ln(k)}{k})^k < e^{k\ln(k)/k} = \frac{1}{k}$, ולכן מתקיים $\mathrm{E}[Y_2] = \frac{m}{k}$. על כן $\mathrm{E}[Y_1 + Y_2] < (n \ln k + m)/k$, ומכאן שיש בחירה ספציפית של U_1 שעבורה מתקבל U שזהו חסם עליון על הגודל שלו.

עתה נראה איך אפשר למצוא באופן דטרמיניסטי (בזמן פולינומי ב־n ו־m) קבוצה U_1 שעבורה מתקיים $Y \leq \mathrm{E}[Y]$ עבור $Y = Y_1 + Y_2$, מה שנותן לנו מיידית אלגוריתם דטרמיניסטי למציאת קבוצה U כנדרש. לשם כך עבור כל $i \in V = \{1, \ldots, n\}$ נגדיר מ"מ אינדיקטור X_i עבור המאורע $i \in U_1$. התוחלת המותנה של Y בהינתן $X_1 = b_1, \ldots, X_i = b_i$ (לכל $0 \leq i \leq n$ וסדרה $b_1, \ldots, b_i$) גם ניתנת לחישוב בזמן פולינומי לפי לינאריות התוחלת. עבור Y_1 פשוט מתקיים $\mathrm{E}[Y_1|X_1 = b_1, \ldots, X_i = b_i] = \sum_{j=1}^{i} b_j + (n-i)\frac{\ln(k)}{k}$, ועבור Y_2, נשים לב שאם $F \subset \{1, \ldots, n\}$ היא קשת של ההיפרגרף, אז אם קיים $j \in F \cap \{1, \ldots, i\}$ שעבורו $X_j = 1$ אז ההסתברות שקשת זו תתרום ל־Y_2 היא אפס, ואם לא קיים j כזה אז ההסתברות היא $(1 - \frac{\ln(k)}{k})^{|F \setminus \{1,\ldots,i\}|}$, שגם ניתנת לחישוב יעיל. על כן ניתן לחשב ביעילות גם את $\mathrm{E}[Y_2|X_1 = b_1, \ldots, X_i = b_i]$, מה שמאפשר לנו להשתמש בשיטת התוחלות המותנות על מנת למצוא את U_1.

דרגה, צביעה, מותן

נתחיל עם הגרף המקרי המוגרל לפי $G(n, \alpha/n)$, עבור α שנבחר בהמשך.

בסופו של דבר נרצה להסיר מהגרף קשתות. על כן ננתח כמה קשתות יהיו בתוך כל קבוצה בת לפחות n/k צמתים. עבור קבוצה A קבועה, מספר הקשתות בתוכה הוא סכום של $\binom{|A|}{2}$ משתנים מקריים ב"ת שכל אחד מהם מקבל 1 בהסתברות α/n ו־0 בהסתברות $1 - \alpha/n$. עבור n גדול דיו התוחלת של מספר הקשתות היא לפחות $\frac{\alpha}{3k^2}n$, ולכן לפי חסם צ'רנוף כפלי (שמופיע בתרגול) עם $\delta = \frac{1}{2}$, ההסתברות שיהיו פחות מ־$\frac{\alpha}{6k^2}n$ קשתות כאלו חסומה ע"י $e^{-\alpha n/24k^2}$. נבחר $\alpha = 24k^2$, ואז בהסתברות $1 - o(1)$ (לפי איחוד מאורעות על לא יותר מ־2^n קבוצות אפשריות) בכל קבוצה A כזו יהיו לפחות $\frac{\alpha}{6k^2}n$ קשתות.

עבור גרף G המקיים את הנ"ל, גם אם נסיר ממנו פחות מ־$\frac{\alpha}{6k^2}n$ קשתות, לא נוכל לצבוע אותו ב־k צבעים, בגלל שעדיין לא תהיה לנו קבוצה חסרת־קשתות בת לפחות n/k קשתות. עתה נשתמש בשאלה "קרבה לדרגה קבועה" עם $\beta = \frac{\alpha}{12k^2}n$ ו־$\gamma = \frac{1}{3}$, כדי להבטיח שבהסתברות לפחות $\frac{2}{3}$ נוכל להסיר לא יותר מ־βn קשתות ולקבל גרף עם דרגה חסומה ע"י d, כאשר d הוא הקבוע המתאים התלוי ב־α, β, γ, שלושה קבועים שנבחרו כאן עם תלות ב־k בלבד.

עתה ננתח את מספר המעגלים מגודל קטן מ־$C \log n$ עבור C כל שהוא. נחסום עבור n גדול דיו את תוחלת מספר המעגלים: $\sum_{i=3}^{C\log n - 1} \frac{n!}{2i(n-i)!} \cdot \frac{\alpha^i}{n^i} \leq \sum_{i=3}^{C\log n - 1} \alpha^i \leq \alpha^{C \log n}$. עבור $C > 0$ קטן מספיק (תלוי ב־α וב־β שתלויים רק ב־k), התוחלת הזו תהיה קטנה ממש מ־$\frac{1}{3}\beta n$, ולכן מאי שוויון מרקוב בסיכוי לפחות $\frac{2}{3}$ יהיו בגרף פחות מ־βn מעגלים, שניתן להסיר את כולם ע"י כך שמסירים קשת אחת מכל מעגל.

מאיחוד מאורעות, בסיכוי חיובי הגרף G יקיים את כל שלושת התנאים: הוא לא יהיה k־צביע כל עוד מסירים ממנו פחות מ־$\frac{\alpha}{6k^2}n = 2\beta n$ קשתות, יהיה ניתן להפוך אותו לבעל דרגה חסומה ע"י d ע"י הסרה של לא יותר מ־βn קשתות, ויהיה ניתן להפוך אותו לחסר מעגלים מגודל קטן מ־$C \log n$ ע"י הסרה של פחות מ־βn קשתות נוספות. לכן, אם לוקחים G כזה ומסירים ממנו את הקשתות עבור סיפוק תנאי הדרגה והמעגלים, מקבלים את הגרף המבוקש.

רוקדים על שתי החתונות

נשתמש בשיטה של הגרלה עם תיקונים (בסגנון דומה לזה של השאלה "קבוצות ב"ת בהיפרגרפים"). נגדיר תהליך מקרי בן מספר שלבים לבנית הקבוצה A שתמיד תקיים את הנדרש מבחינת השכנים של כל צומת ב־$V \setminus A$, ונחסום את התוחלת של הגודל $|A|$.

- בשלב הראשון, מגדירים קבוצה A_0 ע"י כך שבאופן ב"ת, לכל צומת $v \in V$, נבחר אותו להיות ב־A_0 בהסתברות $2\log(d)/d$, באופן ב"ת לכל הצמתים. נשים לב שמתקיים $\mathrm{E}[|A_0|] = 2n\log(d)/d$.
- בשלב השני, נסמן ב־B_0 את קבוצת הצמתים ב־$V \setminus A_0$ שכל שכניהם ב־$V \setminus A_0$. מלינאריות התוחלת מתקיים $\mathrm{E}[|B_0|] \leq n \cdot (1 - 2\log(d)/d)^{d+1} < ne^{-2\log(d)} < n\log(d)/d$ עבור d גדול דיו. הסבר קצר: לכל צומת $v \in V$ יש לפחות d שכנים, ולכל אחד מהם כולל v עצמו יש הסתברות ב"ת של $1 - 2\log(d)/d$ להשאר ב־$V \setminus A_0$. נגדיר $A_1 = A_0 \cup B_0$. בפרט אין אף צומת ב־$V \setminus A_1$ שכל שכניו נמצאים ב־$V \setminus A_1$.
- בשלב השלישי, נסמן ב־B_1 את קבוצת הצמתים ב־$V \setminus A_1$ שכל שכניהם ב־A_1. על מנת לחסום את התוחלת של B_1, נחשוב על צומת $v \in V$ מדרגה $d' \geq d$. לכל צומת w שהוא שכן של v, יש הסתברות $2\log(d)/d$ להיות ב־A_0, והסתברות חסומה ע"י $\log(d)/d$ להיות ב־B_0 (ראו את החישוב של הסעיף הקודם). על כן תוחלת מספר השכנים של v שנמצאים ב־A_1 חסומה ע"י $3d'\log(d)/d$, ולכן לפי אי־שוויון מרקוב הסיכוי שכל d' השכנים יהיו ב־A_1 (הדבר שיכניס את v ל־B_1) חסום ע"י $3\log(d)/d$. על כן מתקיים $\mathrm{E}[|B_1|] \leq 3n\log(d)/d$. לבסוף נגדיר $A = A_1 \cup B_1$. נשים לב שהפעולה הזו לא השפיעה כלל על מיקום השכנים של הצמתים שנשארו ב־$V \setminus A$ (יחסית למצב שהיה עם A_1), ולכן עדיין לא יהיו צמתים ב־$V \setminus A$ שכל שכניהם ב־$V \setminus A$, ז"א ש־A מקיימת את כל הנדרש בשאלה.

נפעיל שוב את לינאריות התוחלת, ונקבל $\mathrm{E}[|A|] = \mathrm{E}[|A_0|] + \mathrm{E}[|B_0|] + \mathrm{E}[|B_1|] \leq 6n\log(d)/d$. לכן יש בחירה ספציפית של A עם חסם הגודל הזה שמקיימת את הנדרש.

למת הבידוד

זיווגים מתחמקים

נתון גרף (לא מכוון) G עם קבוצת צמתים V וקבוצת קשתות E, ונתונה גם תת־קבוצה $F \subseteq E$ של קשתות "אדומות". הראו איך בונים אלגוריתם ב־RNC (ראו את האלגוריתם לזיווגים מושלמים שמשתמש בלמת הבידוד מתוך חוברת ההרצאות), אשר במידה ויש זיווגים מושלמים בגרף, יחזיר את המספר המינימלי של קשתות אדומות שחייבים להכליל בזיווג מושלם כל שהוא.

הסבר: האלגוריתם יפלוט מספר טבעי או "אין זיווג". בהסתברות לפחות $\frac{1}{2}$ האלגוריתם יפלוט את המספר הנכון. בכל מקרה האלגוריתם לעולם לא יפלוט מספר שהוא קטן יותר מהמספר הנכון (השגיאה היא חד־כיוונית במובן זה שיכול להיות רק שהאלגוריתם יפלוט "אין זיווג" או מספר גבוה מדי), ולעולם לא יפלוט מספר כל שהוא אם כלל אין זיווג.

קיום מסלול בגרף מכוון

עבור גרף מכוון בעל n צמתים G וצמתים s, t נרצה לבדוק האם קיים מסלול מ־s ל־t. הראו קיום רדוקציה הסתברותית (ועם סבוכיות מקום LogSpace) של קלט של הבעיה הכללית G, s, t לקלט G', s', t' בעל הפרמטרים הבאים: אם אין מסלול ב־G מ־s ל־t, אז גם אין אף מסלול ב־G' מ־s' ל־t' (בהסתברות 1), ואם יש מסלול ב־G, אז בהסתברות לפחות $\Omega(n^{-3})$ יש ב־G' מסלול יחידי מ־s' ל־t'.

הערה: רדוקציה זו משמשת בהוכחה של Wigderson לכך שמתקיים $\mathrm{NL/poly} \subseteq \oplus\mathrm{L/poly}$.

בידוד של שני מבנים

נניח ש־A קבוצה בת m איברים, וש־$\mathcal{F}$ היא משפחה של תתי קבוצות של A. הראו שאם מגרילים באופן מקרי ויוניפורמי (וב"ת) פונקציית משקל $w : A \to \{1, \ldots, n\}$, אז בסיכוי $1 - \frac{3m}{n}$ לפחות גם האיבר ב־$\mathcal{F}$ בעל המשקל המינימלי וגם האיבר ב־$\mathcal{F}$ בעל המשקל השני הכי קטן הם יחידים.

בידוד רב־קבוצות

נניח כי $\mathcal{F}$ היא משפחה של רב־קבוצות הנלקחת מהקבוצה $A = \{1, \ldots, n\}$, כשכל איבר מ־A רשאי להופיע עד r פעמים באיבר מ־$\mathcal{F}$. נבחר באופן מקרי ויוניפורמי משקל $w(a) \in \{1, \ldots, c\}$ לכל $a \in A$ ונקבע לכל $F \in \mathcal{F}$ את המשקל המושרה עליו, כלומר $w(F) = \sum_{a \in F} w(a)$ כאשר כל איבר נסכם כמספר מופעיו בקבוצה. הוכיחו כי בהסתברות של לפחות $1 - \frac{rn}{c}$ קיים איבר יחיד ב $\mathcal{F}$ עם משקל מינימום.

כמו כן, הציגו דוגמא למקרה בו ההסתברות לקיום איבר יחיד עם משקל מינימום היא פחות מ־$1 - \frac{n}{c}$ (מספיק למצוא דוגמה עבור n ו־c ספציפיים).

פתרונות לתרגילים על למת הבידוד

זיווגים מתחמקים

האלגוריתם שנבנה יהיה דומה לאלגוריתם מחוברת ההרצאות עבור זיווגים מושלמים, עם ההבדלים הבאים:

- המשקלות w_{ij} עבור $ij \in E$ יוגרלו בדיוק כמו באלגוריתם הזיווג המקורי.
- עבור $ij \notin E$ ועבור $ij \in E \setminus F$, נגדיר את a_{ij} בדיוק כמו באלגוריתם המקורי.
- עבור $ij \in F$, נגדיר $a_{ij} = 2^{w_{ij}+n^3}$ אם $i < j$ ונגדיר $a_{ij} = -2^{w_{ij}+n^3}$ אם $i > j$.
- עבור הפלט, אם $\det A = 0$ האלגוריתם יפלוט "אין זיווג". אחרת, האלגוריתם יחשב את המספר המקסימלי k כך ש־2^k מחלק את $\det A$, ויפלוט את $\lfloor k/2n^3 \rfloor$.

ראשית נשים לב שאם בכלל לא קיימים זיווגים מושלמים, אז תמיד יתקיים $\det A = 0$ בדומה להוכחה המקורית.

עבור המשך הניתוח, לכל זיווג מושלם M נסמן ב־$r(M)$ את מספר הקשתות האדומות (מ־F) שהוא מכיל. נסמן ב־r את מספר הקשתות האדומות המינימלי בכל זיווג מושלם כל שהוא, וב־$\mathcal{F}$ את משפחת הזיווגים המושלמים שעבורם $r(M) = r$. נפעיל את למת הבידוד עבור $\mathcal{F}$ (ולא עבור משפחת כל הזיווגים המושלמים), ונקבל שבהסתברות לפחות $\frac{1}{2}$ קיים זיווג מושלם יחידי $M_0 \in \mathcal{F}$ שעבורו $w(M_0)$ הוא מינימלי.

מעתה נתמקד במקרה שבו יש זיווגים מושלמים בגרף, ובתוכם יש זיווג מושלם M_0 יחיד עם משקל מינמלי, מבין כל הזיווגים המושלמים שלהם מספר קשתות אדומות מינימלי. להשלמת ההוכחה (בדומה להוכחה המקורית עבור זיווגים מושלמים), צריך להראות שבמקרה זה הדטרמיננטה $\det A = \sum_{\sigma \in S_n} (-1)^{\operatorname{sgn}(\sigma)} \prod_{i=1}^{n} a_{i\sigma(i)}$ אינה מתאפסת ואף אינה מתחלקת ב־$2^{2(r+1)n^3}$, וכמו כן צריך להראות שבכל מקרה (גם כאשר לא קיים M_0 יחיד כמתואר) הדטרמיננטה תתחלק ב־2^{2rn^3}. הניתוח של איברי הסכום נעשה באופן הבא:

עבור פרמוטציות שמכילות עגילים מגודל אי־זוגי, יהיה קיזוז הדדי בדיוק כמו בהוכחה המקורית. עבור פרמוטציה σ שניתנת לפירוק לזיווגים מושלמים M_1ו־M_2 (כזכור כל פרמוטציה ללא עגילים אי־זוגיים ניתנת לפירוק מסוג זה), יתקיים $|\prod_i a_{i\sigma(i)}| = 2^{w(M_1)+w(M_2)+(r(M_1)+r(M_2))n^3}$. בפרט אם σ_0 היא הפרמוטציה המתארת מעבר הלוך ושוב על הקשתות של M_0 , אז מתקיים $|\prod_i a_{i\sigma_0(i)}| = 2^{2w(M_0)+2rn^3}$.

אם מתקיים $M_1 \notin \mathcal{F}$ ו/או $M_2 \notin \mathcal{F}$, אז מכיוון שלכל זיווג M מתקיים $n/2 \leq w(M) \leq n^3/2$, יתקיים $w(M_1) + w(M_2) + (r(M_1) + r(M_2))n^3 \geq w(M_1) + w(M_2) + (2r + 1)n^3 > 2w(M_0) + 2rn^3$, ולכן

$|\prod_i a_{i\sigma(i)}|$ יתחלק ב־$2^{2w(M_0)+2rn^3+1}$. אם כן מתקיים $M_1, M_2 \in \mathcal{F}$ אבל לא שניהם זהים ל־M_0, אז בדומה להוכחה בחוברת גם כאן נקבל ש־$|\prod_i a_{i\sigma(i)}|$ יתחלק ב־$2^{2w(M_0)+2rn^3+1}$. בזאת כיסינו את כל המקרים עבור כל הפרמוטציות השונות מ־σ_0, ולכן בסכום של $\det A$ יהיה איבר יחידי שאינו מתחלק ב־$2^{2w(M_0)+2rn^3+1}$, ז"א שהדטרמיננטה (הסכום הכולל) בפרט אינה מתאפסת ואינה מתחלקת ב־$2^{2w(M_0)+2rn^3+1}$. מכיוון שמתקיים $n/2 \le w(M_0) \le n^3/2$, נובע מכך שהדטרמיננטה אינה מתחלקת ב־$2^{2(r+1)n^3}$.

מצד שני, כל האיברים בסכום שמגדיר את הדטרמיננטה, גם במקרה שלא קיים M_0 יחיד כמתואר (כאשר יכולים להיות זיווגים אחרים עם r קשתות אדומות שמשיגים את מינימום המשקל), תמיד יתחלקו ב־$2^{w(M_1)+w(M_1)+2rn^3}$ (כאשר M_1 ו־M_2 הם זיווגים המרכיבים את הפרמוטציה המתאימה לאיבר), ובפרט יתחלקו ב־2^{2rn^3}. מאלו נובע שהערך k שהאלגוריתם מחשב (במקרה שלא פולט "אין זיווג") יקיים $2rn^3 \le k < 2(r+1)n^3$ כאשר M_0 הוא יחיד, ובכל מקרה יתקיים $2rn^3 \le k$. מכך נובע ש־$\lfloor k/2n^3 \rfloor$ יהיה זהה ל־r כאשר M_0 יחיד (דבר הקורה בהסתברות $\frac{1}{2}$ לפחות), ובכל מקרה לא יהיה קטן מ־r.

קיום מסלול בגרף מכוון

אם נגריל לכל קשת ב־G משקל שנבחר יוניפורמית ובאופן ב"ת מהתחום $\{1,\dots,2n^2\}$, אז במידה ויש בגרף מסלול כל שהוא מ־s ל־t, לפי למת הבידוד בהסתברות של לפחות $\frac{1}{2}$ יהיה עתה מסלול יחידי עבורו סכום המשקלות הוא מינימלי; מכיוון שמסלול מינימלי הוא בהכרח פשוט, אפשר לראות כל מסלול כתת קבוצה של הקשתות. בנוסף, המשקל של המסלול המינימלי בוודאי לא יעלה על n^3. את הרדוקציה מ־G לגרף החדש G' נבצע עתה באופן הבא.

ראשית, נבחר באופן יוניפורמי מספר $1 \le l \le n^3$. לכל צומת v בגרף המקורי, יהיו בגרף החדש $l+1$ צמתים שיסומנו $(v,0),\dots,(v,l)$. עתה לכל קשת u,v בגרף המקורי נבחר יוניפורמית את המשקל $w(u,v)$ שלה מתוך $\{1,\dots,2n^2\}$, ונצרף לגרף החדש את כל הקשתות מהטיפוס $(u,i),(v,i+w(u,v))$ עבור $0 \le i \le l-w(u,v)$. הרדוקציה היא ב־LogSpace בגלל שאפשר "לשכוח" את $w(u,v)$ מייד לאחר כתיבת הקשתות המתאימות לה ב־G', ולשחרר את הזיכרון עבור משקל הקשת הבאה (אגב, בהרבה מודלים חשוביים מתירים לאלגוריתם עם מקום מוגבל לקבל מראש רשימה של כל הטלות המטבע שלא על חשבון הזיכרון שלו, אולם זה לא היה נוסח השאלה כאן).

עתה נבחן מה הסיכוי שיש ב־G' מסלול יחידי מ־$(s,0)$ ל־(t,l). אם בגרף G אין מסלול מ־s ל־t, אז לא קשה לראות שאין בגרף החדש כל מסלול מ־$(s,0)$ ל־(t,l). מצד שני, אם יש בגרף G מסלול כזה, אז מספר המסלולים בגרף החדש זהה למספר המסלולים ב־G שעבורם סכום המשקלות הוא בדיוק l (כולל מסלולים לא פשוטים). אם ב־G היה מסלול, אז בהסתברות של לפחות $\frac{1}{2}$ יהיה מסלול יחידי בעל משקל מינימלי. במידה וזה אכן קרה, ההסתברות ש־l יהיה שווה למשקל המסלול המינימלי היחיד היא n^{-3}. לכן בהסתברות $\frac{1}{2}n^{-3} = \Omega(n^{-3})$ שני המאורעות יקרו, ובמצב זה יהיה ב־G' מסלול יחידי מ־$(s,0)$ ל־(t,l).

בידוד של שני מבנים

ההוכחה נעשית בדומה להוכחה של למת הבידוד המקורית. עבור פונקציית המשקל שנבחרה w, נגדיר לכל $a \in A$ את הערכים הבאים: W_a יהיה המשקל המינימלי מבין כל איברי $\mathcal{F}$ המכילים את a. $\overline{W}_a$ יהיה המשקל המינימלי מבין אלו שאינם מכילים את a. W'_a יהיה המשקל של האיבר בעל המשקל השני הכי קל מאלו שמכילים את a. $\overline{W}'_a$ יהיה המשקל של האיבר השני הכי קל מאלו שאינם מכילים את a.

נקרא לאיבר a "חד משמעי ביותר" אם גם $W_a \ne \overline{W}_a$ וגם $W_a \ne \overline{W}'_a$ וגם $W'_a \ne \overline{W}_a$. בדומה להוכחת הלמה המקורית, הערכים $\overline{W}_a$, $\overline{W}'_a$, $W_a - w(a)$ ו־$W'_a - w(a)$ כולם אינם תלוים בערך $w(a)$, אלא רק בערכים של w עבור האיברים ב־$A \setminus \{a\}$. על כן כל אחד מאי השוויונים הרצויים מתקיים בהסתברות לפחות $1 - \frac{1}{n}$, ומכאן שכולם יתקיימו בהסתברות לפחות $1 - \frac{3}{n}$. לכן (שוב ע"י שימוש בחסם על איחוד מאורעות) בהסתברות לפחות $1 - \frac{3m}{n}$ כל איברי A הם חד משמעיים ביותר. עתה כל שנותר להוכיח הוא שבהינתן שכל איברי A מקיימים זאת, גם האיבר של $\mathcal{F}$ בעל המשקל המינימלי וגם האיבר בעל המשקל השני הכי קטן הם יחידים.

האיבר בעל המשקל המינימלי הוא יחיד מכיוון שע"פ ההנחה, כל איברי A הם בפרט חד משמעיים במובן של ההוכחה המקורית של למת הבידוד. נסמן איבר זה ב־F. עתה, נניח בסתירה שיש שני איברים $G_1, G_2 \in \mathcal{F}$ בעלי אותו משקל שהוא השני הכי קטן ב־$\mathcal{F}$. בלי הגבלת הכלליות, נניח שקיים איבר $a \in A$ השייך ל־G_1 ואינו שייך ל־G_2. עתה קיימים שני מקרים.

אם $a \in F$, אז $W'_a = w(G_1)$, מכיוון שמבין כל איברי $\mathcal{F}$ המכילים את a האיבר F הוא (היחיד) בעל המשקל המינימלי ו־G_1 הוא בעל המשקל השני הכי קטן. בנוסף לכך, $\overline{W}_a = w(G_2) = w(G_1)$, מכיוון שמבין האיברים שאינם מכילים את a האיבר G_2 יהיה בעל המשקל המינימלי (שהרי F אינו נמצא שם). בזאת קיבלנו סתירה ל־$W'_a \neq \overline{W}_a$. באותו האופן, עבור המקרה $a \notin F$ נקבל סתירה ל־$W_a \neq \overline{W}'_a$, ושני המקרים ביחד מסיימים את ההוכחה.

בידוד רב־קבוצות

ההוכחה דומה להוכחה של למת הבידוד הרגילה, רק שכאן נצטרך לשמור $r+1$ ערכים לכל $a \in A$ במקום שניים. נסמן $W_{0,a}, \ldots, W_{r,a}$ כאשר $W_{s,a}$ הוא משקל הרב־קבוצה מ־$\mathcal{F}$ בעלת משקל המינימום מבין אלו המכילות את a בדיוק s פעמים.

נאמר ש־a רב־משמעי אם קיימים $s < t$ כך ש־$W_{s,a} = W_{t,a}$. נניח כי ישנן שתי רב־קבוצות $F_1, F_2 \in \mathcal{F}$ כך שמשקלן מינימלי. קיים איבר a שנמצא ב־F_1 בדיוק s פעמים וב־F_2 מספר פעמים אחר t. כך מתקיים עבור הקבוצות $W_{t,a} = w(F_1) = w(F_2) = W_{s,a}$, כלומר a רב־משמעי. כעת נסמן $V_{s,a} = W_{s,a} - s \cdot w(a)$. ערך זה נקבע לחלוטין על ידי משקלי חברי $A \backslash \{a\}$. כעת, $W_{s,a} = W_{t,a}$ אם ורק אם $V_{t,a} - V_{s,a} = (s-t)\, w(a)$, וכמו בהרצאה, זה קורה בערך אחד של $w(a)$ לכל היותר ולכן מתרחש בהסתברות של $\frac{1}{c}$ לכל היותר.

מחסם האיחוד על פני s, t, a נקבל שההסתברות לקיום איבר בעל משקל מינימלי יחיד ב־$\mathcal{F}$ היא לפחות $1 - \frac{n}{c}\binom{r+1}{2}$, אבל אנחנו רוצים חסם חזק יותר. לשם כך אנחנו נקבע פונקציית משקל w על $A \backslash \{a\}$ ונראה כי למעשה ישנם לכל היותר r ערכים אפשריים ל־$w(a)$ שיהפכו את a לרב־משמעי.

נקבע את המשקלות כאמור, ונאמר כי i מרבה את s עבור a אם קיים $t > s$ כך שקביעת $w(a) = i$ גורמת לכך ש־$W_{s,a} = W_{t,a}$, ושניהם מינימליים מבין $W_{0,a}, \ldots, W_{r,a}$. נראה כי לכל s יש לכל היותר i אחד שמרבה אותו עבור a (ולכן אין יותר מ־r ערכי j שגורמים לריבוי כל שהוא). נניח על דרך השלילה כי $j > i$ שניהם מרבים את s עבור a. נקבע את $W_{k,a}(i)$ להיות $W_{k,a}$ כאשר אנחנו בוחרים $w(a) = i$. עבור $w(a) = j$ נקבל $W_{k,a}(j) = W_{k,a}(i) + (j-i)\, k$, כי משקל יתר הקבוצה נשאר זהה, ורק הוספנו $j - i$ למשקל של a, שיש לו k מופעים. לכן לכל $t' > s$ אנחנו מקבלים

$$
\begin{aligned}
W_{t',a}(j) - W_{s,a}(j) &= W_{t',a}(i) - W_{s,a}(i) + \left(t' - s\right)(j-i) \\
&> W_{t',a}(i) - W_{s,a}(i) \geq 0
\end{aligned}
$$

כשהאי שוויון האחרון הוא מכיוון ש־i מרבה את s ולכן $W_{s,a}(i)$ הוא מינימלי מבין ה־$W_{k,a}(i)$. לכן אין אף t' שיכול לגרום לכך ש־j ירבה את s עבור a.

כעת נראה את הדוגמה שמראה שלא נוכל לשפר זאת ל $1 - \frac{n}{c}$. הדוגמא תהיה עם $n = 2, c = 3$ וההסתברות תהיה 0. נראה משפחה של רב־קבוצות מעל $A = \{a, b\}$ עבורה כל פונקציית משקל $w : A \to \{1, 2, 3\}$ תתן שתי רב־קבוצות ממשקל זהה. נסמן ב (i, j) את הרב־קבוצה המכילה i מופעים של a ו־j מופעים של b. המשפחה שנגדיר היא $\mathcal{F} = \{(13,0), (10,1), (8,2), (5,4), (4,5), (2,8), (1,10), (0,13)\}$. אפשר לוודא את התכונה על פני מעבר על פני תשע פונקציות המשקל האפשריות.

המומנט השני

הפרדה ע"י פונקציה לינארית

תהי $A \subseteq \{0,1\}^n$ קבוצה כל שהיא בת k איברים בקוביה הבוליאנית ה־n מימדית. הראו שקיימת פונקציה לינארית $f : (\mathbb{Z}_2)^n \to \mathbb{Z}_2$ כך שמתקיים עבור מספר האיברים ב־A המאפסים את הפונקציה הבוליאנית f אי השוויון $|\{v \in A | f(v) = 0\}| = \frac{1}{2}k \pm O(\sqrt{k})$.

תתי גרפים של גרפים צפופים

נתון גרף G כל שהוא אשר מספר הקשתות שלו הוא αn^2 עבור $0 \le \alpha < \frac{1}{2}$ מתאים. נגריל תת גרף מושרה מקרי H, ע"י כך שכל צומת של G תיבחר להיות צומת של H בהסתברות $\frac{1}{2}$, באופן ב"ת בצמתים האחרים. הראו שבסיכוי $1 - o(1)$ מספר הקשתות ב־H הוא $\frac{1}{4}\alpha n^2 + o(n^2)$.

סף למעגלים

הראו, עבור כל k קבוע, שהפונקציה $f(n) = 1/n$ היא פונקצית סף לקיום מעגל מגודל k בגרף מקרי: אם $p(n) = o(1/n)$, אז הגרף המקרי $G(n, p(n))$ מכיל מעגל מגודל k בסיכוי $o(1)$, ואם $p(n) = \omega(1/n)$, אז הגרף המקרי $G(n, p(n))$ מכיל מעגל מגודל k בסיכוי $1 - o(1)$.

הבהרות: המעגלים לא חייבים להיות תתי־גרף מושרים. המשמעות של "לכל k קבוע" היא שלחסם ההסתברות מותר להיות תלוי ב־k. הוא צריך לשאוף לערך האסימפטוטי לכל k "בנפרד" (לא במידה שווה) כאשר n שואף ל־∞.

הדרכה: סימונים מופשטים מתאימים יכולים לעזור הרבה. למשל, אפשר לסמן ב־$\mathcal{C}$ את קבוצת כל המעגלים מאורך k בגרף השלם עם n צמתים, ועבור מעגל $C \in \mathcal{C}$ לסמן ב־X_C את מ"מ האינדיקטור עבור המאורע ש־C מעגל בגרף המקרי G. כדאי גם לשים לב שקבוצה בת $l < k$ קשתות בתוך מעגל C מגודל k מכסה לפחות $l+1$ צמתים.

קרובים למטרה

נתונים $a_1, \ldots, a_n$, כולם מספרים ממשיים בין 0 ו־1, שמתקיים עבורם $\sum_{i=1}^n (a_i)^2 = 1$. הראו שקיים קבוע $\alpha > 0$ שאינו תלוי ב־n או בערכי a_i, כך שאם מגרילים את $b_1, \ldots, b_n \in \{-1, 1\}$ באופן יוניפורמי וב"ת, אז בהסתברות לפחות α מתקיים $|\sum_{i=1}^n b_i \cdot a_i| \le 1$.

רמז: זה עוזר לפצל למקרים לפי מהו ערך ה־a_i המקסימלי.

הערה: אפשר להשתמש בטענת השאלה על מנת להוכיח את הטענה הגיאומטרית הבאה: לכל על־מישור שעובר דרך הראשית ב־$\mathbb{R}^n$ (על־מישור כזה הוא קבוצת נקודות המאופיינת כקבוצת הפתרונות של משוואה לינארית בודדת), לפחות $\alpha 2^n$ מנקודות הקבוצה $\{-1,1\}^n \subset \mathbb{R}^n$ הם במרחק שאינו עולה על 1 מעל־המישור. אתם מוזמנים לחשוב איך עושים את זה.

פסוקיות לא מאוד מסופקות

הראו שלכל ϵ קיים קבוע C, כך שלכל נוסחת 3-NAE-SAT עם n משתנים ו־$m > Cn$ פסוקיות קיימת הצבה המספקת לפחות $(\frac{3}{4} - \epsilon)m$ ולא יותר מ־$(\frac{3}{4} + \epsilon)m$ מהפסוקיות. פסוקית של 3-NAE-SAT מסתפקת אם לא כל שלושת הליטרלים מקבלים אותו ערך (ז"א אם לא כולם אפס ולא כולם אחד), וההנחה היא שכל פסוקית תלויה בדיוק בשלושה משתנים שונים.

פרישה נרחבת

הראו שקיים קבוע $c>0$ עם התכונה הבאה: לכל מספר ראשוני p, כל $k>0$, וכל תת־קבוצה A של השדה $\mathbb{Z}_p$ (שדה השלמים מודולו p) המקיימת $|A|=k$, קיים איבר $x\in\mathbb{Z}_p$ (שונה מ־0) כך שהקבוצה $x\cdot A=\{xa:a\in A\}$ מכילה איבר מכל מקטע ב־$\mathbb{Z}_p$ שגודלו הוא לפחות $cp/\sqrt{k}$. מקטע מאורך l הוא קבוצה מהצורה $I_{r,l}=\{r,\ldots,r+l-1\}$ עבור $r\in\mathbb{Z}_p$, והחשבון הוא ב־$\mathbb{Z}_p$, כך שלמשל הקבוצה $I_{-\lfloor l/2\rfloor,l}=\{p-\lfloor l/2\rfloor,\ldots,p-1,0,\ldots,\lceil l/2\rceil-1\}$ היא דוגמה לקטע כזה.

הדרכה: כדאי לצמצם את מספר הקטעים שצריך להראות שייכות אליהם. כדאי גם "להוסיף פרמטר מיותר" ולהראות קודם שקיימים x ו־y כך שהקבוצה $x\cdot A+y=\{xa+y:a\in A\}$ מקיימת את התכונה הדרושה.

ריכוז במרחב

נביט במרחב הוקטורי $\mathbb{Z}_p^n$ (מעל השדה $\mathbb{Z}_p$) עבור $p\geq 3$ ראשוני ו־$n\geq 2$. נתונה קבוצה $A\subseteq\mathbb{Z}_p^n\setminus\{0\}$ כך ש־$|A|=\frac{p^n-1}{2}$. נבחר תת מרחב U ממימד 2 יוניפורמית. הראו, עבור p גדל ובאופן ב"ת ב־n (ז"א עם התכנסות במידה שווה ביחס ל־n), כי בהסתברות $1-o(1)$ מתקיים שגודל החיתוך $A\cap U$ הוא בין $\left(\frac{1}{2}+o(1)\right)(p^2-1)$ לבין $\left(\frac{1}{2}-o(1)\right)(p^2-1)$.

פתרונות לתרגילים על המומנט השני

הפרדה ע"י פונקציה לינארית

אנו נגריל את f באופן מקרי. כזכור, כל פונקציה לינארית מ־$(\mathbb{Z}_2)^n$ ל־$\mathbb{Z}_2$ נתונה ע"י ווקטור $u\in\mathbb{Z}_2$, כך שלכל $v\in(\mathbb{Z}_2)^n$ מתקיים $f(v)=u\cdot v$ (הכפל הוא כפל ווקטורי מעל $\mathbb{Z}_2$). אנו נגריל את u באופן יוניפורמי (כל קורדינטה באופן ב"ת באחרות).

לכל $v\in A$ נסמן ב־X_v את משתנה האינדיקטור עבור המאורע ש־$f(v)=0$. מתקיים $\mathrm{E}[X_v]=\frac{1}{2}$, וכן לכל $v\neq w$ המ"מ X_v ו־X_w הם בלתי תלויים בזוגות. להוכחת הטענה השניה נניח בלי הגבלת הכלליות שמתקיים $v_n=0$ ו־$w_n=1$: ניתן להניח בה"כ שקיימת קורדינטה i שמתאפסת ב־v ולא ב־w (אחרת נחליף את v ו־w), והטיעון הוא זהה אם $i=n$. עתה נחשב (למשל) את $\Pr[X_v=1\wedge X_w=1]$ באמצעות נוסחת ההסתברות השלמה באופן הבא:

$$
\begin{aligned}
\Pr[X_v=1\wedge X_w=1] &= \sum_{\substack{\alpha_1,\ldots,\alpha_{n-1}\\ v\cdot(\alpha_1,\ldots,\alpha_{n-1},0)=1}} \Pr[u_1=\alpha_1,\ldots,u_{n-1}=\alpha_{n-1}]\Pr[w\cdot(\alpha_1,\ldots,\alpha_{n-1},u_n)=1]\\
&= \sum_{\substack{\alpha_1,\ldots,\alpha_{n-1}\\ v\cdot(\alpha_1,\ldots,\alpha_{n-1},0)=1}} 2^{1-n}\cdot\frac{1}{2}=2^{n-2}\cdot 2^{1-n}\cdot\frac{1}{2}=\frac{1}{4}
\end{aligned}
$$

מכך נובע שהמ"מ המתאר את מספר איברי A המאפסים את f, הנתון ע"י $X=\sum_{v\in A}X_v$, מקיים $\mathrm{E}[X]=\frac{k}{2}$ וכן $\mathrm{V}[X]=\frac{k}{4}$. ממשפט צ'בישף נובע עתה שבהסתברות גדולה מ־0 יתקיים (למשל) $|X-\frac{k}{2}|\leq\sqrt{k}$, ולכן קיימת פונקציה f המקיימת את המבוקש.

תתי גרפים של גרפים צפופים

נסמן את קבוצת הקשתות של G ב־E, ולכל $e\in E$ נגדיר את המשתנה X_e להיות שווה ל־1 אם הקשת e מוכלת בתת הגרף H שנבחר (ז"א ששני הצמתים שלה נבחרו כצמתים של H), ושווה ל־0 אחרת. לבסוף

נגדיר את $X = \sum_{e\in E} X_e$, ונשים לב שמ"מ זה זהה בערכו למספר הקשתות של H. מלינאריות התוחלת, $\mathrm{E}[X] = \sum_{e\in E} \mathrm{E}[X_e] = \frac{1}{4}\alpha n^2$. עתה נחסום את $\mathrm{V}[X]$.

עתה נרשום $\mathrm{V}[X] = \sum_{e,f\in E} \mathrm{Cov}[X_e, X_f]$. אם e ו־f זרות צמתים אז $\mathrm{Cov}[X_e, X_f] = 0$, ואחרת עדיין מתקיים $\mathrm{Cov}[X_e, X_f] < 1$ (לא צריך כאן חסם יותר טוב). סה"כ קיבלנו $\mathrm{V}[X] < 2|E|n \cdot 1 \le n^3$. ממשפט צ'בישף נובע עתה שלכל η מתקיים $\Pr\left[|X - \frac{1}{4}\alpha n^2| > \eta\right] < n^3/\eta^2$. עבור $\eta = n^{7/4}$ נקבל שבסיכוי $1 - n^{-1/2} = 1 - o(1)$ מתקיים $X = \frac{1}{4}\alpha n^2 \pm n^{7/4} = \frac{1}{4}\alpha n^2 + o(n^2)$ כנדרש.

סף למעגלים

למען פישוט הסימונים, נסמן p במקום $p(n)$. כמו כן נשתמש בסימון $\mathcal{C}$ עבור קבוצת כל המעגלים מגודל k בגרף השלם עם n צמתים (זוהי "קבוצת כל המעגלים האפשריים ב־G"), כפי שנעשה בהדרכה לשאלה.

ראשית נוכיח במהירות את החסם התחתון: עבור מעגל ספציפי C מגודל k, הסיכוי שהוא יהיה מוכל ב־G הוא בדיוק p^k. מספר המעגלים האפשריים הוא $|\mathcal{C}| = O(n^k)$; לאלו הזוכרים את בניית הגרפים עם מותן גבוהה מההרצאה על הגרלה עם תיקונים, המספר המדויק הוא $\frac{k!}{2k}\binom{n}{k}$. מכאן, לפי החסם על איחוד מאורעות, הסיכוי לקיום מעגל מתוך $\mathcal{C}$ ב־G חסום ע"י $O(n^k p^k) = O((np)^k)$. מכיוון ש־k קבוע, אם $p = o(1/n)$ אז הביטוי הזה שואף ל־0 כאשר n שואף ל־∞.

עתה נעבור לחסם העליון. נניח אם כן שמתקיים $p = \omega(1/n)$. לכל $C \in \mathcal{C}$ נסמן ב־X_C את מ"מ האינדיקטור עבור הקיום של C ב־G, ונסמן את מספר המעגלים הכולל ב־G ב־$X = \sum_{C\in\mathcal{C}} X_C$. בדומה למה שנעשה בהרצאה עבור פונקצית סף לקיום K_4, נראה אצלנו שעבור k קבוע מתקיים $\mathrm{V}[X] = o((\mathrm{E}[X])^2)$, אשר לפי אי־שוויון צ'בישף ייתן לנו את החסם המבוקש על ההסתברות לאי־קיום מעגל. קודם כל נשים לב שחישוב מהיר נותן $\mathrm{E}[X] = |\mathcal{C}|p^k = \Omega((np)^k)$.

עתה, עבור $0 \le l \le k$, נסמן ב־$\mathcal{I}_l$ את קבוצת זוגות המעגלים (C, C') שמקיימים $C \neq C'$ ויש להם בדיוק l צמתים משותפים (הדרישה $C \neq C'$ היא על מנת שלא להכליל זוגות מהצורה (C, C) ב־$\mathcal{I}_k$). מתקיים אם כן $\mathrm{V}[X] = \sum_{C,C'\in\mathcal{C}} \mathrm{Cov}[X_C, X_{C'}] = \sum_{C\in\mathcal{C}} \mathrm{V}[X_C] + \sum_{l=0}^{k} \sum_{(C,C')\in\mathcal{I}_l} \mathrm{Cov}[X_C, X_{C'}]$. עתה נחסום כל אחד מהסכומים הנ"ל לחוד.

- חישוב ישיר עבור סכום השונויות נותן $\sum_{C\in\mathcal{C}} \mathrm{V}[X_C] = |\mathcal{C}|(p^k - p^{2k}) = O(n^k p^k) = O((np)^k)$.
- עבור $(C, C') \in \mathcal{I}_0 \cup \mathcal{I}_1$ יש אי־תלות של X_C ב־$X_{C'}$, ולכן הסכומים עבור $l = 0, 1$ מתאפסים.
- עבור $1 < l \le k$, נשים לב שאם $(C, C') \in \mathcal{I}_l$, אז יש ל־$C$ ו־C' לכל היותר $l - 1$ קשתות משותפות (זה נובע מהטענה שהוזכרה בהדרכה לשאלה). על כן $\mathrm{Cov}[X_C, X_{C'}] \le \Pr[C, C' \subset G] \le p^{2k+1-l}$ (בדומה לחסימת הקוואריאנס עבור מ"מ אינדיקטור מההרצאה). כמו כן מתקיים $|\mathcal{I}_l| = O(n^{2k-l})$, מכיוון שאפשר להגדיר כל זוג מעגלים כזה ע"י סדרה של $2k - l$ צמתים (יכולות להיות מספר סדרות שמגדירות את אותו הזוג $(C, C') \in \mathcal{I}_l$, אבל בכל מקרה אנחנו מעונינים בחסם עליון). מכל אלו מתקיים $\sum_{(C,C')\in\mathcal{I}_l} \mathrm{Cov}[X_C, X_{C'}] = O(n^{2k-l}p^{2k+1-l}) = O(p(np)^{2k-l}) = O((np)^{2k-l})$.

לסיום, נשווה כל אחד מהמחוברים (יש לנו $k - 1$ מחוברים שלא מתאפסים) לחוד מול $(\mathrm{E}[X])^2$. מכיוון ש־np שואף ל־∞ (כזכור הנחנו שמתקיים $p = \omega(1/n)$), מתקיים $(np)^k = o((np)^{2k}) = o((\mathrm{E}[X])^2)$, ובאופן דומה לכל $1 < l \le k$ מתקיים $(np)^{2k-l} = o((\mathrm{E}[X])^2)$. לכן זה נכון גם עבור הסכום של המחוברים (שמספרם קבוע), ז"א שמתקיים $\mathrm{V}[X] = o((\mathrm{E}[X])^2)$.

קרובים למטרה

נסמן מ"מ $X = \sum_{i=1}^{n} b_i \cdot a_i$ (שתלוי בהגרלה של $b_1, \ldots, b_n$). אנחנו נרצה לתת חסם תחתון גדול מאפס עבור המאורע $|X| \le 1$, אבל אי אפשר לעשות את זה ע"י שימוש ישיר באי־שוויון צ'בישף (או משפט ריכוז אחר), כי

מתקיים $\mathrm{V}[X]=\sum_{i=1}^{n}(a_i)^2=1$. בלי הגבלת הכלליות, נניח שמתקיים $a_n=\max\{a_1,\ldots,a_n\}$, ונחלק למקרים לפי גודל a_n. נשים לב שתמיד מתקיים $a_n\leq 1$.

מקרה ראשון: $a_n\geq\frac{1}{2}$. במקרה זה נסמן מ"מ של הסכום ללא האיבר האחרון $Y=\sum_{i=1}^{n-1}b_i\cdot a_i$. מתקיים אז $\mathrm{V}[Y]=\sum_{i=1}^{n-1}(a_i)^2\leq\frac{3}{4}$ (וגם $\mathrm{E}[Y]=0$), ולכן לפי אי־שוויון צ'בישף מתקיים $|Y|\leq\frac{3}{2}$ בהסתברות לפחות $\frac{2}{3}=1-\frac{3/4}{(3/2)^2}$. נגדיר B לפי $B=1$ אם $Y\geq 0$ ו־$B=-1$ אם $Y<0$. נשים לב שאם מתקיים גם $|Y|\leq\frac{3}{2}$ וגם $b_n=-B$ אז מתקיים $\big||Y|-a_n\big|\leq 1$ $=|Y+b_n\cdot a_n|=|X|$. זאת מכיוון שאם $|Y|\leq 1$ אז מתקיים $\big||Y|-a_n\big|\leq\max\{|Y|,a_n\}\leq 1$, ואם $1<|Y|\leq\frac{3}{2}$ אז מתקיים $\big||Y|-a_n\big|=|Y|-a_n\leq\frac{3}{2}-\frac{1}{2}=1$.

על מנת לסיים במקרה הזה, נשים לב שמכיוון ש־b_n מוגרל באופן ב"ת ב־$b_1,\ldots,b_{n-1}$, הוא יהיה ב"ת בערך של Y (שתלוי רק ב־$b_1,\ldots,b_{n-1}$) ובפרט גם ב־B. על כן, בהסתברות לפחות $\frac{2}{3}\cdot\frac{1}{2}=\frac{1}{3}$ יתקיים גם $|Y|\leq\frac{3}{2}$ וגם $b_n=-B$, ובפרט יתקיים אז $|X|\leq 1$.

מקרה שני: $a_n<\frac{1}{2}$, ואז גם $a_i<\frac{1}{2}$ לכל $1\leq i\leq n$. נסמן ב־j את האינדכס המינימלי שעבורו $\sum_{i=1}^{j}(a_i)^2\geq\frac{3}{8}$. מכיוון שמתקיים $\sum_{i=1}^{j-1}(a_i)^2<\frac{3}{8}$, מ־$a_i<\frac{1}{2}$ נובע $\sum_{i=1}^{j}(a_i)^2<\frac{5}{8}$. בנוסף מתקיים $\sum_{i=j+1}^{n}(a_i)^2\leq 1-\frac{3}{8}=\frac{5}{8}$. נגדיר עתה שני מ"מ, $Y=\sum_{i=1}^{j}b_i\cdot a_i$ ו־$Z=\sum_{i=j+1}^{n}b_i\cdot a_i$, כך שמתקיים $X=Y+Z$. כמו כן נגדיר $B=1$ אם $Y\geq 0$ ו־$B=-1$ אם $Y<0$, וכן $C=1$ אם $Z\geq 0$ ו־$C=-1$ אם $Z<0$. הדבר לשים לב הוא ש־Y (ו־B התלוי רק בו) הוא ב"ת ב־Z (ו־C התלוי רק בו), כי $b_1,\ldots,b_j$ נבחרים באופן ב"ת לחלוטין ב־$b_{j+1},\ldots,b_n$.

עתה נשתמש באי־שוויון צ'בישף עם $\mathrm{V}[Y]\leq\frac{5}{8}$ ו־$\mathrm{E}[Y]=0$ על מנת להסיק שמתקיים $|Y|\leq 1$ בהסתברות לפחות $\frac{3}{8}=1-\frac{5}{8}$, ובאופן דומה נסיק שמתקיים $|Z|\leq 1$ בהסתברות לפחות $\frac{3}{8}$. המאורעות האלו הם ב"ת, ולכן שניהם מתקיימים בהסתברות לפחות $\frac{9}{64}$. במידה ושני המאורעות מתקיימים, אם $|Y|=0$ ו/או $|Z|=0$ אז בטוח יתקיים $|X|=|Y+Z|\leq 1$. במידה ושני אלו שונים מאפס, נשים לב שהמאורע $B=-C$ מתקיים בהסתברות בדיוק $\frac{1}{2}$, מכיוון שלכל $\alpha_1,\ldots,\alpha_n\in\{-1,1\}$, ההסתברות עבור המאורע $b_1=\alpha_1,\ldots,b_n=\alpha_n$ שווה להסתברות עבור המאורע $b_1=\alpha_1,\ldots,b_j=\alpha_j,b_{j+1}=-\alpha_{j+1},\ldots,b_n=-\alpha_n$ (שני המאורעות קורים בהסתברות 2^{-n} בדיוק). כמו כן, כאשר $B=-C$ וגם $|Y|,|Z|\leq 1$, מתקיים $|X|\leq\max\{|Y|,|Z|\}\leq 1$.

בסה"כ, במקרה השני יש לנו חסם תחתון של לפחות $\frac{1}{2}\cdot\frac{9}{64}=\frac{9}{128}$ על ההסתברות עבור $|X|\leq 1$. יחד עם המקרה הראשון, כיסינו את כל המקרים האפשריים עבור $a_1,\ldots,a_n$ המקיימים את תנאי השאלה.

פסוקיות לא מאוד מסופקות

אנו נגריל הצבה לקבוצת המשתנים $x_1,\ldots,x_n$ של המשתנים באופן יוניפורמי וב"ת. נסמן ב־X_i את משתנה האינדיקטור עבור המאורע "הפסוקית ה־i הסתפקה", וב־$X=\sum_{i=1}^{m}X_i$ את המ"מ של מספר הפסוקיות שהסתפקו. באופן דומה למה שחושב עבור SAT בכתה מתקיים $\mathrm{E}[X]=\frac{3}{4}m$, ומייד נוכיח עבור בחירה מתאימה של C שיתקיים $\mathrm{V}[X]\leq\frac{1}{4}\epsilon^2m^2$. מכאן ינבע לפי חוק צ'בישף שבסיכוי לפחות $\frac{3}{4}$ ההצבה שלנו תהיה כנדרש.

ראשית נחשב את $\mathrm{Cov}[X_i,X_j]$ לפי ניתוח למקרים:

- אם לפסוקיות ה־i וה־j אין יותר ממשתנה אחד משותף, אז שני מאורעות ההסתפקות הם בלתי תלויים זה בזה (בגלל שידיעת ערך של משתנה אחד אינה משנה את סיכוי ההסתפקות של פסוקית NAE), ולכן מתקיים $\mathrm{Cov}[X_i,X_j]=0$.
- אם לשתי הפסוקיות יש שלושה משתנים משותפים וזו אינה אותה פסוקית אז הקווריאנס אינו חיובי, $\mathrm{Cov}[X_i,X_j]\leq 0$. אם זוהי אותה פסוקית, ז"א ש־$i=j$, אז הקווריאנס זהה לשונות של X_i, שהיא קטנה מ־1 (הערה: היפוך שלושת הליטרלים מביא אותנו למצב של "אותה פסוקית", אם כי זה לא משנה הרבה את החישוב אם אנו מרשים כאלו כפילויות גם).
- אם יש שני משתנים משותפים בדיוק, אז למרות שניתן לחסום באופן יותר מדוייק נסתפק בחסם הפשוט

$\text{Cov}[X_i,X_j]=\text{E}[X_iX_j]-(\frac{3}{4})^2<1$. מספר הזוגות של פסוקיות כאלו חסום ע"י $24m(n-3)<24mn$ (עבור פסוקית מסויימת יש 3 בחירות של זוג משתנים להחתך בהם איתה, ולכל אחת מהן 4 אפשרויות לסימנים עבורם, $n-3$ דרכים לבחור את האיבר הנוסף בפסוקית שחותכת אותה, ושתי אפשרויות לסימן שלו).

עתה ניתן לחסום את השונות של X:

$$\text{V}[X]=\sum_{1\le i,j\le m}\text{Cov}[X_i,X_j]<0+m+24mn<25mn$$

לסיום ההוכחה, בוחרים $C=100\epsilon^{-2}$, על מנת שיתקיים $\frac{1}{4}\epsilon^2m^2>\frac{1}{4}\epsilon^2Cnm=25mn$.

פרישה נרחבת

אם מוכיחים עבור $x,y\in\mathbb{Z}_p$ ש־$x\cdot A+y$ חותך כל קטע מהצורה $I_{r,l}$ עבור $l\ge cp/\sqrt{k}$, אז הדבר יהיה נכון גם עבור $x\cdot A$, מכיוון שמתקיים $(x\cdot A)\cap I_{r,l}=\emptyset$ אם ורק אם $(x\cdot A+y)\cap I_{r+y,l}=\emptyset$. עתה נגריל את $x,y\in\mathbb{Z}_p$ באופן יוניפורמי וב"ת, ונראה שבהסתברות חיובית $x\cdot A+y$ תקיים את התכונה המבוקשת (בהגרלה אנחנו מרשים ל־x לקבל ערך 0 בשביל הניתוח ההסתברותי, בכל מקרה זה לא יהיה הערך שיקיים בסוף את תוצאת השאלה). את הערך של c נבחר בהמשך. כמו כן נניח ש־p גבוה מספיק (יספיק למשל להניח $p\ge 100$), כי אפשר יהיה (עם הגדלה אפשרית בערך של c) לוודא שתוצאת השאלה תהיה טריביאלית עבור p קטנים מדי.

נסמן $m=\lfloor cp/2\sqrt{k}\rfloor$, ולכל $0\le i<\lceil p/m\rceil<4\sqrt{k}/c$ נגדיר את $J_i=I_{i\cdot m,m}$. אם $x\cdot A+y$ חותך את כל הקטעים $J_0,\ldots,J_{\lceil p/m\rceil-1}$ אז הוא יחתוך כל קטע $I_{r,l}$ לכל $0\le r<p$ ולכל $l\ge cp/\sqrt{k}$, פשוט כי מתקיים אז $J_{\lceil r/m\rceil}\subset I_{r,l}$.

נבחן עתה את ההסתברות ש־$x\cdot A+y$ אינו חותך את J_i עבור i קבוע כל שהוא. לכל $a\in A$ נגדיר את X_a להיות משתנה האינדיקטור עבור המאורע "$x\cdot a+y\in J_i$", וכן נגדיר את $X=\sum_{a\in A}X_a$. המאורע ש־$x\cdot A+y$ אינו חותך את J_i הוא בדיוק המאורע "$X=0$". חישוב ישיר שמשתמש בלינאריות התוחלת נותן לנו $\text{E}[X]=|A|\cdot m/p>\frac{1}{4}c\sqrt{k}$.

עבור חישוב המומנט השני, נראה שלכל $a\neq b$ מתקיים ש־X_a ו־X_b הם בלתי־תלויים. עבור $u,v\in\mathbb{Z}_p$ כל שהם (שווים או שונים), יש פתרון יחיד עבור x ו־y למערכת המשוואות $x\cdot a+y=u\ \wedge\ x\cdot b+y=v$, כי המדובר במערכת לא־מנוונת. על כן ההסתברות עבור המאורע "$x\cdot a+y\in J_i\ \wedge\ x\cdot b+y\in J_i$" היא בדיוק m^2/p^2, ז"א שהמאורעות (ולכן המשתנים המתאימים להם) הם ב"ת.

מזאת נובע שמתקיים $\text{V}[X]=\sum_{a\in A}\text{V}[X_a]<|A|\cdot m/p<\frac{1}{2}c\sqrt{k}$. משני אלו, לפי אי־שוויון צ'בישף, $\Pr[X=0]\le\text{V}[X]/(\text{E}[X])^2<8/c\sqrt{k}$. בחירה של $c=6$ תוודא שמתקיים $(4\sqrt{k}/c)\cdot(8/c\sqrt{k})<1$, ובמקרה כזה לפי איחוד מאורעות יש סיכוי חיובי לכך שכל הקטעים $J_0,\ldots,J_{\lceil p/m\rceil-1}$ יחתכו ע"י $x\cdot A+y$. על כן בפרט יש בחירה ספציפית של x ו־y שתתן לנו את המבוקש.

ריכוז במרחב

מסתכלים על תת המרחב כתוצאה מבחירה יוניפורמית של שני ווקטורים ב"ת $u,v\in\mathbb{Z}_p^n$ ומעבר לתת המרחב הנפרש (מספר הבסיסים הפורשים זהה לכל תת מרחב אפשרי ממימד 2, ולכן זו תהיה בחירה יוניפורמית של תת המרחב), ואז עבור כל α,β שאינם שניהם 0 מגדירים את $X_{\alpha,\beta}$ כמשתנה האינדיקטור עבור המאורע "$\alpha u+\beta v\in A$". נשים לב שגודל החיתוך של תת המרחב עם A נתון ע"י $X=\sum_{\alpha,\beta\in\mathbb{Z}_p}X_{\alpha,\beta}$, וכן שמתקיים $\text{E}[X_{\alpha,\beta}]=\frac{1}{2}$ לכל α,β ולכן $\text{E}[X]=\frac{1}{2}(p^2-1)$.

אם (α,β) ו־(α',β') הם ב"ת כזוג ווקטורים ב־$\mathbb{Z}_p^2$, אז עבור זוג משתני האינדיקטור המתאימים מתקיים $\mathrm{Cov}[X_{\alpha,\beta},X_{\alpha',\beta'}] = \mathrm{E}[X_{\alpha,\beta}\cdot X_{\alpha',\beta'}]-\frac{1}{4} \le \frac{1}{2}\cdot(\frac{p^n-1}{2(p^n-p)})-\frac{1}{4} = \frac{p-1}{4(p^n-1)} \le \frac{p-1}{4(p^2-1)}$. הפיתוח הוא באמצעות חסימת $\Pr[X_{\alpha',\beta'}=1|X_{\alpha,\beta}=1]$ באמצעות חישוב ישיר (אפילו אם אנחנו יודעים בדיוק את $\alpha u+\beta v$, יש p^n-p אפשרויות שוות הסתברות עבור $\alpha' u+\beta' v$). מספר הזוגות של וקטורים ב"ת ב־$\mathbb{Z}_p^2$ חסום ע"י מספר הזוגות הכולל של וקטורים שונים מ־0, שזה $(p^2-1)^2$.

אם (α,β) ו־(α',β') אינם ב"ת עדיין מתקיים $\mathrm{Cov}[X_{\alpha,\beta},X_{\alpha',\beta'}]\le\frac{1}{2}-\frac{1}{4}=\frac{1}{4}$, ויש לא יותר מ־$(p-1)(p^2-1)$ זוגות כאלה. סה"כ מתקיים $\mathrm{V}[X]=\sum_{\alpha,\alpha',\beta,\beta'}\mathrm{Cov}[X_{\alpha,\beta},X_{\alpha',\beta'}]\le(\frac{1}{4}+\frac{1}{4})(p-1)(p^2-1)<(p^2-1)^{3/2}$. עתה אפשר לסיים ע"י שימוש באי שוויון צ'בישף: $\Pr[|X-\frac{1}{2}(p^2-1)|>(p^2-1)^{7/8}]<(p^2-1)^{-1/4}=o(1)$, כאשר בפרט $(p^2-1)^{7/8}=o(p^2-1)$ עבור p גדל, וביטויי החסם אינם תלויים ב־n.

חסימת סטיות גדולות

קירוב להסתברות

צריכים לקרב התפלגות לא ידועה μ מעל $\{1,\ldots,n\}$. חשבו על הדרך הבאה: לוקחים q דגימות ב"ת $i_1,\ldots,i_q$ כך שכל i_j מוגרל לפי ההתפלגות μ באופן ב"ת בדגימות האחרות. מגדירים את ההתפלגות ν לפי הנוסחה $\Pr_\nu[i]=|\{j:i_j=i\}|/q$, ז"א שמקרבים את ההסתברות ל־i לפי מספר הפעמים היחסי שתוצאה זו הופיעה ב־q הדגימות שלקחנו. הראו שלכל ϵ קיים C, כך שאפשר בצורה זו עם $q=Cn$ דגימות לוודא שמתקיים $d(\mu,\nu)\le\epsilon$ (המדובר במרחק ה־Variation Distance מהפרק הראשון של חוברת זו) בהסתברות $1-o(1)$.

מתמקדים

נתון תהליך הסתברותי שפולט ערכים ב־$\mathbb{R}$. לא נתון כלום על ההתפלגות של הערכים, פרט לזה שכל הרצה של התהליך היא ב"ת בכל ההרצות הקודמות, וכן שקיימים מספרים $a<b$, כך שבהסתברות לפחות $\frac{2}{3}$ מתקבל ערך בין a ל־b. הראו, לכל $0<\delta<1$, שבאמצעות $O(\log(1/\delta))$ קריאות לתהליך אפשר לקבל ערך שבהסתברות לפחות $1-\delta$ יהיה בין a ל־b.

הבהרה: אין מידע מהם a ו־b, פרט לזה שהם קיימים. על האלגוריתם שאתם בונים לפלוט את הערך המבוקש ללא מידע מוקדם על a ו־b.

חיפוש בינארי עם מעט שקרים

נזכיר את האלגוריתם (הדטרמיניסטי) לחיפוש איבר נתון ברשימה ממוינת בת n איברים באמצעות $\lceil\log_2 n\rceil$ השוואות: מתחילים מהתחום $\{1,\ldots,n\}$. בכל שלב משווים את האיבר הנתון עם האיבר האמצעי בתת הרשימה המתאימה לתחום, ובהתאם עוברים לתת־תחום שגודלו כחצי מגודל התחום בסוף השלב הקודם.

עתה נניח שאנו רוצים לחפש איבר נתון ברשימה ממוינת, אולם כל פעם שאנו משווים את האיבר הנתון עם איבר ברשימה, בסיכוי של 1% נקבל את התשובה ההפוכה לאמת. ליתר דיוק: אין לנו יכולת לקרוא את האיברים מהרשימה אלא רק להשוות אותם. אם תוצאת ההשוואה היא "$>$" אז בסיכוי 1% נקבל את התשובה "$<$", ואם תוצאת ההשוואה היא "$<$" אז בסיכוי 1% נקבל את התשובה "$>$". לא יהיו תשובות שגויות אף פעם ביחס ל־"$=$".

כתבו אלגוריתם שמוצא את האיבר ברשימה הממוינת, אשר רץ בתוחלת זמן שהיא עדיין $O(\log n)$. מותר להניח שהאיבר הנתון אכן קיים ברשימה, ויש להקפיד שניתוח זמן הריצה אכן יהיה נכון ביחס לתוחלת (לא רק "בהסתברות גבוהה הזמן הוא קצר").

מחלקות סיבוכיות

המחלקה BPP מוגדרת כמחלקת השפות שעבורן קיים אלגוריתם הסתברותי אשר רץ בזמן פולינומי, ונותן את התשובה הנכונה בהסתברות $\frac{2}{3}$. המחלקה $\mathrm{P/poly}$ מוגדרת כמחלקת השפות כך שלכל n קיים אלגוריתם דטרמיניסטי אשר רץ בזמן פולינומי ונותן את התשובה הנכונה לכל קלט מאורך n: שימו לב שבניגוד ל־P, זה אינו חייב להיות אותו אלגוריתם ל־n שונים; רק חסם הזמן הפולינומי חייב להיות אחיד לכל ה־n, וכן הוא חייב לחסום את אורך התיאור של האלגוריתם הספציפי ל־n. הוכיחו שמתקיים $\mathrm{BPP} \subseteq \mathrm{P/poly}$.

להקיא את החוכמה

נתונה לנו המשפחה $\mathcal{F}$ של תתי־קבוצה של $\{1, \ldots, n\}$ כמו בשאלה "לבלוע את החוכמה" מהפרק על לינאריות התוחלת, וכן מרחבי ההסתברות μ ו־ν שמוגדרים בשאלה ההיא. הראו שאם מתקיים $pn \geq q$, אז מתקיים $\Pr_\nu[A] > (1 - e^{-(p-q/n)^2 n/2p})\Pr_\mu[A]$.

צריכים לחסוך

נתונה פונקציה $f : \{0,1\}^n \to \{0,1,*\}$, כאשר ערך של $f(x) = *$ פירושו הוא "כל תשובה עבור x היא נכונה". נניח שיש אלגוריתם הסתברותי שקורא את הקלט $x \in \{0,1\}^n$ ב־q מקומות לכל היותר, ולכל x שעבורו $f(x) \in \{0,1\}$, בהסתברות לפחות $\frac{2}{3}$ פולט את הערך הנכון עבור $f(x)$ (כאשר $f(x) = *$ אז מותר לאלגוריתם לפלוט כל ערך בכל הסתברות). הראו שיש אלגוריתם כזה (גם עם הסתברות לפחות $\frac{2}{3}$ לתשובה נכונה כאשר $f(x) \in \{0,1\}$) שקורא את x ב־$3q$ מקומות לכל היותר, וכל ההסתברות שלו מבוססת על הטלה של $k = \log(n) + O(1)$ מטבעות לכל היותר. באופן פורמלי האלגוריתם החדש מקבל ערך $r \in \{0,1\}^k$ שמוגרל יוניפורמית מתוך 2^k האפשרויות, ובהסתמך עליו ועל הקריאות המבוצעות מהקלט הוא פועל דטרמיניסטית.

הערה קטנה: הסיבה שמאפשרים ערכי "$*$" ב־f היא שיש מעט פונקציות "מעניינות" $f : \{0,1\}^n \to \{0,1\}$ שעבורם ידועים אלגוריתמים שנותנים תשובה נכונה בהסתברות לפחות $\frac{2}{3}$ ללא קריאה של $\Omega(n)$ מקומות מהקלט (אין לכך השפעה מהותית על ההוכחה הנדרשת).

הבהרה גדולה: אין שום דרישות על סיבוכיות החישוב (או אפילו אם האלגוריתם ניתן לחישוב). מבחינתכם יכול להיות שפשוט יש לאלגוריתם החדש רשימה עצומה של 2^k אלגוריתמים דטרמיניסטים, והוא בוחר להריץ אחד מתוכם לפי r.

תלויים באופן טוב

נתונים לנו משתנים $X_1, \ldots, X_m$, שכל אחד מהם מקבל ערכים ב־$\{-1,1\}$, אבל הם יכולים להיות תלויים זה בזה. נתון רק שלכל $1 \leq k \leq m$ מתקיים $\Pr[X_k = 1 | X_1 = a_1, \ldots, X_{k-1} = a_{k-1}] \leq \frac{1}{2}$, לכל סדרת ערכים אפשרית $a_1, \ldots, a_{k-1}$. הוכיחו באופן פורמלי שגם כאן מתקיים $\Pr[X > a] < e^{-a^2/2m}$, כאשר מסמנים $X = \sum_{i=1}^m X_i$.

לפספס את הרכבת

כזכור, בשאלה "לתפוס את הרכבת" מהפרק על בניה וניתוח של מרחבי הסתברות, היה נתון שההסתברות שיהיה i שעבורו $X_i = 1$ קטנה ממש מ־1. עתה עליכם להוכיח את הטענה הזו.

הערה: מותר (ורצוי) להשתמש בסימונים של השאלה "לתפוס את הרכבת", וכן של הפתרון הרשמי שלה. עם זאת צריך לזכור שההוכחה שם מסתמכת על התוצאה שאתם צריכים להוכיח כאן, ולכן צריך להזהר מטיעונים מעגליים.

תתי־קבוצה מקריים

נניח שאנו מגרילים $2n$ תתי־קבוצה מגודל 3 של $\{1, \ldots, n\}$, כשכל קבוצה לכשעצמה מוגרלת יוניפורמית מכל תתי־הקבוצה האפשריים באופן ב"ת בקבוצות האחרות. הראו שבסיכוי $1 - e^{-\Theta(n)}$ ניתן לבחור n קבוצות מתוכן

כך שאף איבר של $\{1,\ldots,n\}$ לא יופיע ביותר מ־40 קבוצות שונות.

שוויון הנפש

הראו שלכל $0<\alpha<2$ הדבר הבא מתקיים לכל מספר טבעי n גדול מ־N_α מתאים, ולכל $k\leq \alpha n/\log(n)$ (הלוגריתם הוא בבסיס 2): אם $A_1,\ldots,A_k$ הם תתי־קבוצה כל שהם של $\{1,\ldots,n\}$, אז קיימים שני תתי־קבוצה $B_1,B_2\subseteq\{1,\ldots,n\}$, שונים זה מזה, שעבורם מתקיים $|A_i\cap B_1|=|A_i\cap B_2|$ לכל $1\leq i\leq k$.

רמז: לא צריך לנתח הגרלה במקביל של שני תתי־קבוצה. ניתוח הסתברותי של האפשרויות עבור החיתוכים עם תת־קבוצה אחת יעזור.

קליקים בממוצע

נבחר גרף לפי $G\left(n,\frac{1}{2}\right)$. הראו כי לכל k קבוע בהסתברות של לכל היותר $e^{-\Theta(n)}$ מספר ה־k־קליקים בגרף יהיה יותר מ־$\binom{n}{k}\cdot 2^{-\binom{k}{2}+1}$ או פחות מ־$\binom{n}{k}\cdot 2^{-\binom{k}{2}-1}$.

פתרונות לתרגילים על חסימת סטיות גדולות

קירוב להסתברות

אנחנו נשתמש בתוצאת סעיף ב של השאלה "קרבה בין התפלגויות" מהפרק הראשון של חוברת זו. אנחנו נראה, עבור בחירה מתאימה של C, שבהסתברות $1-o(1)$ לכל מאורע E יתקיים $|\Pr_\mu[E]-\Pr_\nu[E]|\leq\epsilon$, ומכאן נובע ש־$d(\mu,\nu)\leq\epsilon$, מכיוון שתמיד קיים מאורע E שעבורו ההפרש $|\Pr_\mu[E]-\Pr_\nu[E]|$ שווה בדיוק ל־$d(\mu,\nu)$.

נניח שאנחנו לוקחים את הדגימות $i_1,\ldots,i_q$ לפי המתואר בשאלה, ונחסום עבור מאורע E ספציפי את ההסתברות (ביחס לתהליך לקיחת הדגימות) שמתקיים $|\Pr_\mu[E]-\Pr_\nu[E]|\leq\epsilon$. כזכור, מאורע במרחב הסתברות מתואר ע"י תת־קבוצה של קבוצת הבסיס שלו, ובמקרה שלנו E היא תת־קבוצה של $\{1,\ldots,n\}$. לפי ההגדרה של ν, מתקיים $\Pr_\nu[E]=|\{1\leq j\leq q:i_j\in E\}|/q$.

עבור התהליך של לקיחת הדגימות נגדיר מ"מ $X_1,\ldots,X_q$ כאשר X_j הוא האינדיקטור של המאורע "$i_j\in E$", ונגדיר את $X=\sum_{j=1}^q X_j$. בפרט מתקיים $\Pr_\nu[E]=X/q$. הדבר הבא לשים לב הוא ש־$X_1,\ldots,X_q$ הם ב"ת לחלוטין זה בזה, וכל X_j מקבל 1 בהסתברות $\Pr_\mu[E]$ בדיוק, כך שמתקיים $\Pr_\mu[E]=\mathrm{E}[X]/q$.

עבור $q=Cn$, יתקיים $|\Pr_\mu[E]-\Pr_\nu[E]|>\epsilon$ אם ורק אם $X>\mathrm{E}[X]+\epsilon Cn$ או $X<\mathrm{E}[X]-\epsilon Cn$. לפי החסם על סטיות גדולות מחוברת ההרצאות, ההסתברות לאיחוד שני המאורעות האלו חסומה ע"י $2\exp(-2\epsilon^2Cn)$. אם למשל נבחר $C=\lceil 1/\epsilon^2\rceil$ אז הסתברות זו תהיה $o(2^{-n})$.

לבסוף, נשים לב שיש בדיוק 2^n אפשרויות עבור המאורע E (היינו יכולים לצמצם ל־2^{n-1} אפשרויות אם היינו שמים לב לכך שאי השוויון מתקיים עבור E אם ורק אם הוא מתקיים עבור $\neg E$). על כן, לפי איחוד מאורעות, בהסתברות $1-o(1)$ לא יהיה מאורע E שעבורו $|\Pr_\mu[E]-\Pr_\nu[E]|>\epsilon$, כפי שרצינו להוכיח.

מתמקדים

נניח שמתקיים $\delta<\frac{1}{3}$, אחרת אפשר פשוט לקחת הרצה בודדת של התהליך ההסתברותי ולפלוט את התוצאה. עבור k אי־זוגי שנבחר בקרוב, נבצע k פעמים את התהליך ההסתברותי, ומתוך סדרת התוצאות $c_1,\ldots,c_k$ נפלוט את החציון, ז"א את הערך c_j שעבורו $|\{i:c_i\leq c_j\}|\geq\frac{k}{2}$ וגם $|\{i:c_i\geq c_j\}|\geq\frac{k}{2}$ (צורת הכתיבה כאן מתאימה גם למקרה שיש שוויונות בסדרת הערכים).

על מנת שהפלט יהיה קטן מ־a, צריך שיהיו לפחות $\frac{k}{2}$ קריאות לתהליך ההסתברותי שנתנו ערך קטן מ־a, ועל מנת שהפלט יהיה גדול מ־b, צריך שיהיו לפחות $\frac{k}{2}$ קריאות שנתנו ערך גדול מ־b. בשני המקרים זה אומר שלפחות $\frac{k}{2}$

מהקריאות לא נתנו ערך בין a ל־b. כמו כן, בשאלה נתון שבכל קריאה בודדת הסיכוי לקבל ערך שאינו בין a ל־b חסום ע"י $\frac{1}{3}$.

נגדיר עתה מ"מ אינדיקטור X_i עבור המאורע "לא מתקיים $a \leq c_i \leq b$". לפי נתוני השאלה $X_1, \ldots, X_n$ הם מ"מ ב"ת שכל אחד מהם מקבל 1 בהסתברות שאינה עולה על $\frac{1}{3}$, ומקבל 0 אחרת. מהניתוח למעלה עולה שבפרט מתקיים $\sum_{i=1}^{k} X_i \geq \frac{k}{2}$ בכל מקרה שפלט האלגוריתם אינו בין a ל־b. לפי חסימת סטיות גדולות (אי־השוויון השני מחוברת ההרצאות), ההסתברות עבור פלט שגוי חסומה ע"י $e^{-2(k/2-k/3)^2/k} = e^{-k/18}$. אם רוצים שחסם זה יהיה קטן מ־δ, אפשר לקחת למשל $k = 18\lceil\log(1/\delta)\rceil + 1 = O(\log(1/\delta))$.

חיפוש בינארי עם מעט שקרים

נבצע כאן גרסא של אלגוריתם החיפוש הבינארי הרגיל, אולם עם תוספת אפשרות של "חזרה לאחור": בשלב הראשון נתחיל עם הקטע $\{1, \ldots, n\}$ כמו באלגוריתם הרגיל. עתה בכל שלב, כאשר בידינו הקטע $\{a, \ldots, b\}$, נבצע השוואה עם איבר הרשימה במקום ה־$\lfloor \frac{a+b}{2} \rfloor$ כמו באלגוריתם הרגיל, אולם בנוסף לכך נבצע השוואה גם עם המקום ה־a וגם עם המקום ה־b. אם קיבלנו שוויון, נעצור כמו באלגוריתם הרגיל (זכרו את ההנחה בשאלה שאף פעם אין תוצאה שקרית ביחס לשוויון). אם תוצאות שלושת ההשוואות מתאימות להנחה שהאיבר נמצא בתחום המקומות $\{a, \ldots, \lfloor \frac{a+b}{2} \rfloor\}$ או נמצא בתחום המקומות $\{\lfloor \frac{a+b}{2} \rfloor + 1, \ldots, b\}$, אז נעבור לחצי הקטע המתאים כפי שהדבר נעשה באלגוריתם החיפוש הבינארי. לבסוף, אם תוצאות ההשוואות מראות שהאיבר המבוקש אינו נמצא כלל בתחום $\{a, \ldots, b\}$, או ששלושת התוצאות אינן קונסיסטנטיות, אז "נחזור לאחור": במקרה זה אנו נגדיל את הקטע חזרה למה שהיה לפני ההקטנה האחרונה (בכל שלב אנו נשמור את ההסטוריה של כל הקטעים עד הקטע הנוכחי שלא חזרנו מהם). אם מצב זה קורה עבור הקטע $\{1, \ldots, n\}$ אז אי אפשר לחזור לאחור, ואז פשוט נשאיר אותו כמות שהוא לאיטרציה הבאה.

עתה ננתח את תוחלת זמן הריצה. אנו נקרא לצעד של האלגוריתם "נכון" אם קרה אחד משני הדברים הבאים: או שהאיבר המבוקש אכן נמצא בתחום המקומות $\{a, \ldots, b\}$ ואנו אכן חצינו את הקטע לתת הקטע המכיל את האיבר, או שהאיבר אינו נמצא בתחום $\{a, \ldots, b\}$ ואנו אכן בצענו צעד לאחור. לכל אפשרות אחרת נקרא "צעד לא נכון". לשם פשטות הניתוח, כאשר האלגוריתם מוצא את האיבר ועוצר, אנו נניח שהוא ממשיך לבצע "צעדים נכונים" בהסתברות 1.

שימו לב שבכל שלב של האלגוריתם, ההסתברות לצעד נכון היא לפחות 97% (ההסתברות הספציפית יכולה להיות תלויה בצעדים קודמים, אבל לא החסם). בנוסף לכך, ברגע שההפרש בין מספר הצעדים הנכונים ומספר הצעדים הלא־נכונים עולה על $\lceil \log_2 n \rceil$ לטובת הנכונים הרי שהאלגוריתם עצר בהצלחה (יתכן כי הוא כבר עצר בהצלחה קודם לכן). לפי חסימת סטיות גדולות, הסיכוי שזה לא קרה עבור $t \geq 10 \log_2 n$ צעדים (שימו לב שתוחלת ההפרש בין הצעדים הנכונים ללא־נכונים היא לפחות $\frac{94}{100}t$) הוא לא יותר מ־$e^{-t/10}$ (בעצם הרבה פחות אבל התעלמנו כאן מקבועים מדויקים): מביטים בסדרת משתנים מקריים המקבלים 1 בצעד נכון ו־-1 בצעד לא נכון, וחוסמים בעזרת חסם סטיות גדולות מתאים. נסמן לבסוף ב־T את המ"מ שמקבל את מספר הצעדים שלקח לאלגוריתם לעצור, ונקבל:

$$
\begin{aligned}
\mathrm{E}[T] &= \sum_{t=1}^{\infty} t\Pr[T=t] = \sum_{j=1}^{\infty}\sum_{i=1}^{j} \Pr\left[T=j\right] = \sum_{t=1}^{\infty}\Pr[T \geq t] \\
&\leq \lceil 10\log_2 n \rceil + \sum_{t=\lceil 10\log_2 n \rceil}^{\infty} \Pr[T \geq t] \\
&\leq \lceil 10\log_2 n \rceil + \sum_{t=\lceil 10\log_2 n \rceil}^{\infty} e^{-t/10} = O(\log_2 n)
\end{aligned}
$$

מחלקות סיבוכיות

נניח ש־L היא שפה השייכת ל־BPP, ונוכיח שהיא ב־P/poly. נניח ש־$p(n)$ הוא חסם זמן פולינומי עבור אלגוריתם הסתברותי המכריע את L, וכן נניח שאלגוריתם זה משתמש רק בבחירות יוניפורמיות ב"ת מתוך $\{0,1\}$ ("הטלות מטבע"). לביסוס ההנחה אפשר להשתמש בשאלה "סימולציה של הסתברות" מפרק התרגילים על השיטה הבסיסית. לכל n נבנה אלגוריתם דטרמיניסטי שנותן תשובות נכונות עבור כל המילים מאורך n, כך שגם זמן הריצה וגם אורך התיאור שלו חסומים ע"י החסם הפולינומי $O(np(n))$.

האלגוריתם ההסתברותי המקורי משתמש עבור מילים מאורך n בלא יותר מ־$p(n)$ מטבעות (הגרלות יוניפורמיות ב"ת מ־$\{0,1\}$), ולכן ניתן לתאר אותו ע"י בחירה מקרית יוניפורמית של מחרוזת ב־$\{0,1\}^{p(n)}$, שבהסתמך עליה ועל הקלט האלגוריתם מגיע להכרעה דטרמיניסטית (כל פעם שהאלגוריתם אמור להטיל מטבע, קוראים במקום זאת את הערך הבא מהמחרוזת שהוגרלה).

עתה נגריל באופן ב"ת $100n$ מחרוזות בינאריות מאורך $p(n)$ שנסמנן $\alpha_1,\ldots,\alpha_{100n}$, ונבחן את $100n$ ההרצות האפשריות של האלגוריתם המתקבלות מהן. אם $a \in \{0,1\}^n$ היא מילה בשפה, אז לפי חסימת סטיות גדולות, הסיכוי שהיא תתקבל עבור לא יותר מ־$51n$ מההרצות הנ"ל חסום ע"י 2^{-n-1} (ולמעשה פחות מכך). בדומה לכך, אם $a \in \{0,1\}^n$ אינה ב־L אז הסיכוי שהיא תתקבל עבור לא פחות מ־$49n$ מההרצות חסום ע"י 2^{-n-1}.

מכאן נובע שקיימת בחירה של $\alpha_1,\ldots,\alpha_{100n}$ שעבורה לכל מילה $w \in \{0,1\}^n$, מילה זו מתקבלת ע"י רב ההרצות המתאימות ל־$\alpha_1,\ldots,\alpha_{100n}$ אם ורק אם $w \in L$. עתה אנו יכולים להרכיב את האלגוריתם שלנו עבור מילים מאורך n עבור הבחירה הנ"ל (שתהווה חלק מתיאור האלגוריתם): בהינתן $w \in \{0,1\}^n$, לכל α_i נבצע את ההרצה המתאימה (באופן דטרמיניסטי בהסתמך על a ו־α_i) ונכתוב את התשובה. האלגוריתם שלנו יקבל את w אם ורק אם לפחות $50n$ מההרצות הנ"ל קיבלו את המילה a.

להקיא את החוכמה

נגדיר מרחב הסתברות τ על זוגות $Q, Q' \subseteq \{1,\ldots,n\}$ באופן הבא:

- ראשית נגריל את Q' לפי ν.
- אם מתקיים $|Q'| \leq q$, אז נקבע $Q = Q'$.
- אם מתקיים $|Q'| > q$, אז נגריל את Q יוניפורמית מתתי־הקבוצה של Q' שיש להם q איברים בדיוק.

כמו בפתרון השאלה "לבלוע את החוכמה" מהפרק על לינאריות התוחלת, נסמן ב־A_1 את המאורע "קיימת $F \in \mathcal{F}$ שעבורה מתקיים $F \subseteq Q$", וב־A_2 את המאורע "קיימת $F \in \mathcal{F}$ שעבורה מתקיים $F \subseteq Q'$". בפרט שוב מתקיים $\Pr_\tau[A_2] \geq \Pr_\tau[A_1]$ מכיוון שתמיד מתקיים $Q \subseteq Q'$. הפעם מתקיים $\Pr_\tau[A_2] = \Pr_\nu[A]$, כי Q' מתפלג באופן זהה תחת τ ו־ν.

עתה נסמן ב־B את המאורע "$|Q'| < q$". מתקיים $\Pr_\tau[B] < e^{-(p-q/n)^2n/2p}$ לפי אי־שוויון צ'רנוף האחרון מהפרק על חסימת סטיות גדולות בחוברת התרגול, כאשר מציבים בו $\mu = pn$ ו־$\delta = (pn-q)/pn = (p-q/n)/p$. כמו כן, נימוק מאוד דומה לזה של "לבלוע את החוכמה" מראה שההתפלגות של Q תחת τ בהינתן המאורע $\neg B$ זהה להתפלגות של Q תחת μ.

מאלו נובע שמתקיים $\Pr_\mu[A] = \Pr_\tau[A_1|\neg B] = \Pr_\tau[A_1 \wedge \neg B]/\Pr_\tau[\neg B] < \Pr_\tau[A_1]/(1-e^{-(p-q/n)^2n/2p})$,

ולכן $\Pr_\nu[A] = \Pr_\tau[A_2] \geq \Pr_\tau[A_1] > (1-e^{-(p-q/n)^2n/2p})\Pr_\mu[A]$ כנדרש.

צריכים לחסוך

נראה איך מתרגמים אלגוריתם הסתברותי שמבצע q קריאות (אבל לא מוגבל בכמות האקראיות שלו) לכזה שמבצע $3q$ קריאות ומקיים את חסם האקראיות הדרוש. ראשית דבר נתרגם אותו לאלגוריתם שמבצע $3q$ קריאות, עדיין

לא מציית לחסם אקראיות כל שהוא, אבל נותן את התשובה הנכונה בהסתברות לפחות $\frac{20}{27}$ (במקום $\frac{2}{3}$). המעבר פשוט: מריצים את האלגוריתם המקורי שלוש פעמים, כל פעם באופן ב"ת בפעמים הקודמות (מבצעים את כל ההגרלות באופן ב"ת בהרצות קודמות), ומחזירים את התשובה שלפחות שתיים מההרצות הנ"ל החזירו. אם האלגוריתם המקורי טעה בהסתברות $\alpha \leq \frac{1}{3}$, אז האלגוריתם החדש יטעה בהסתברות $\alpha^3 + 3(1-\alpha)\alpha^2 \leq \frac{7}{27}$.

עתה נראה תרגום של האלגוריתם החדש לאלגוריתם עם אותו מספר שאילתות, שמשתמש ב־$\log(n) + O(1)$ הטלות מטבע בלבד, ונותן את התשובה הנכונה בהסתברות לפחות $\frac{2}{3}$. לשם כך נסתכל על האלגוריתם ההסתברותי שאנחנו הולכים לתרגם באופן הבא: נניח שכל ההגרלות שהאלגוריתם יכול לעשות במהלך ההרצה נעשות מראש, ובמהלך הריצה של האלגוריתם הוא רק "קורא" אותן יחד עם הקריאות שלו מהקלט. במילים אחרות, מסתכלים על האלגוריתם ההסתברותי כמרחב הסתברות מעל אלגוריתמים דטרמיניסטים (שימו לב אבל שהאלגוריתמים הדטרמיניסטים הנ"ל אינם בד"כ בעלי "תיאור קצר", במקרה הכללי אלגוריתם כזה מתואר כעץ החלטות בגובה חסום ע"י $3q$).

נחשוב עתה על קלט ספציפי $x \in \{0,1\}^n$, וננתח מה קורה כאשר מבצעים l הרצות בלתי־תלויות של האלגוריתם ההסתברותי עליו. באופן מדויק, נרצה לדעת את ההסתברות שפחות מ־$\frac{2}{3}l$ מההרצות הנ"ל יתנו תשובה נכונה. לפי חסימת סטיות גדולות (החסם השני מההרצאה), ההסתברות שזה יקרה חסומה ע"י $e^{-2(2/3-20/27)^2 l} < 2^{-l/100}$. על כן, לפי איחוד מאורעות, עבור $l \geq 100n$ יש סיכוי חיובי ש־l ההרצות הנ"ל יקיימו בו זמנית את התנאי שלכל $x \in \{0,1\}^n$ לפחות $\frac{2}{3}l$ מהן נותנות תשובה נכונה.

שימו לב שלכל $1 \leq i \leq l$, מה שאנחנו מכנים "ההרצה ה־i" משתמשת בסדרה אחת של "הגרלות הנעשות מראש" (ראו את הפסקה הלפני־קודמת), ולא מבצעת שום הגרלות נוספות תוך כדי קריאת $x \in \{0,1\}^n$ עצמו. בהינתן ההגרלות הנ"ל זהו אלגוריתם דטרמיניסטי. כדאי בשלב זה לכתוב הבהרה: כאשר אנחנו בוחרים את "ההרצה ה־i", אנחנו קובעים לא רק את ההגרלות שהאלגוריתם ביצע בפועל לפי הערכים שהוא קרא מ־x מסויים, אלא קובעים את כל ההגרלות שהאלגוריתם היה יכול לבצע בהינתן כל התשובות האפשריות לקריאות שהוא ביצע מהקלט. את האלגוריתם הדטרמיניסטי המתקבל ניתן לכתוב כעץ החלטות שבו כל צומת פנימי מבצע קריאה למקום מסויים בקלט, ושני ילדיו מתארים את המשך ההרצה לפי הערך שנקרא.

נבחר את ה־k המינימלי שעבורו $2^k \geq 100n$ ונבחר $l = 2^k$. שימו לב שבפרט מתקיים $k = \log(n) + O(1)$. עתה נרשום סדרה ספציפית של l אלגוריתמים דטרמיניסטים שמקיימת שלכל $x \in \{0,1\}^n$, לפחות $\frac{2}{3}l$ מתוכם נותנים תשובה נכונה. סדרה כזו קיימת לפי הפסקה הקודמת: כזכור כל הרצה של אלגוריתם הסתברותי ניתנת לתיאור כבחירה הסתברותית של אלגוריתם דטרמיניסטי שאותו מריצים, ומכיוון שיש חסם הסתברות חיובי על כך שלכל $x \in \{0,1\}^n$ לפחות $\frac{2}{3}l$ מההרצות מהן נותנות תשובה נכונה, יש בחירה ספציפית של l הרצות שמקיימות את המבוקש.

לאחר שרשמנו את הרשימה הארוכה הנ"ל, אפשר לנסח את האלגוריתם הסופי שלנו: האלגוריתם יגריל יוניפורמית ערך i מתוך $\{1,\ldots,l\}$, שאפשר ע"י ייצוג בינארי לחשוב עליו כעל התוצאה של הגרלה יוניפורמית מתוך $\{0,1\}^k$, ואז יבצע מרשימת ההרצות את ההרצה ה־i.

תלויים באופן טוב

אפשרות אחת היא לבצע מחדש הוכחה דומה לזו של אי השוויון על חסימת סטיות גדולות, כשמוסיפים בה טיעון של "תלות בכל ערך אפשרי", בדומה להוכחה של אי־שוויון אזומה בפרק על מרטינגלים. להרחבת האופקים נראה כאן טיעון המבוסס על שיטת צימוד.

נבנה כאן סדרה שניה של משתנים $Y_1,\ldots,Y_m$, תלויים ב־$X_1,\ldots,X_m$, שתקיים את הדברים הבאים: לכל i מתקיים $X_i \leq Y_i$ (בהסתברות 1), וכן $Y_1,\ldots,Y_m$ בלתי תלויים לחלוטין זה בזה. משני אלו, בהסתברות גדולה מ־$1-e^{-a^2/2m}$, מתקיים $\sum_{i=1}^m X_i \leq \sum_{i=1}^m Y_i \leq a$.

לאחר שערכי $X_1,\ldots,X_m$ הוגרלו, נגריל את ערכי ה־Y_i באינדוקציה: בהינתן $Y_1,\ldots,Y_{i-1}$, אם $X_i = 1$ אז נקבע

$Y_i = 1$ בהסתברות 1. אם $X_i = -1$, אז נחשב את $\alpha_i = \Pr[X_i = 1|X_1, \ldots, X_{i-1}]$ לפי הערכים שכבר הגרלנו ל־$X_1, \ldots, X_{i-1}$. עתה נבחר $Y_i = 1$ בהסתברות $(\frac{1}{2} - \alpha_i)/(1 - \alpha_i)$ (לפי נתוני השאלה תמיד $\alpha_i \leq \frac{1}{2}$), ונבחר $Y_i = -1$ בהסתברות $\frac{1}{2}/(1 - \alpha_i)$. מההגדרה ברור מיידית שתמיד יתקיים $X_i \leq Y_i$. חישוב מיידי יראה גם שלכל סדרת ערכים של $X_1, \ldots, X_{i-1}$ ושל $Y_1, \ldots, Y_{i-1}$, מתקיים $\Pr[Y_i = 1|Y_1, \ldots, Y_{i-1}, X_1, \ldots, X_{i-1}] = \frac{1}{2}$. לכן $\Pr[Y_i = 1|Y_1, \ldots, Y_{i-1}] = \frac{1}{2}$, ומכאן אפשר להראות באינדוקציה ש־$Y_1, \ldots, Y_m$ בלתי־תלויים (כל סדרת ערכים אפשרית תתקבל בהסתברות 2^{-m} בדיוק).

לפספס את הרכבת

נשים לב שאפשר לתת הגדרה אלטרנטיבית למרחב ההסתברות מעל ערכי הסדרה $X_0, X_1, \ldots$ באופן הבא: נגריל סדרה של מ"מ ב"ת (לחלוטין) $Z_1, Z_2, \ldots$, כאשר כל Z_i יקבל את הערך 1 בהסתברות $\frac{1}{3}$ ויקבל את הערך 0 בהסתברות $\frac{2}{3}$, ולכל $i \geq 0$ נגדיר $X_i = -i + 2\sum_{j=1}^{i} Z_j$. בצורה זו נקבל בדיוק את אותה ההתפלגות על המ"מ $X_0, X_1, \ldots$ כמקודם.

עתה לכל $k \in \mathbb{N}$ נגדיר את המאורע A_k, שאומר שקיים i שעבורו $X_i = k$. בפרט A_1 הוא המאורע "A" מהפתרון הרשמי של "לתפוס את הרכבת", שזה המאורע שאת ההסתברות שלו אנחנו צריכים לחסום כאן. לפני שנמשיך, נשים לב שעל מנת לפתור את השאלה מספיק להראות שקיים k שעבורו $\Pr[A_k] < 1$. יש שתי דרכים להצדיק את זה: האפשרות הראשונה היא להוכיח (בדומה למה שנעשה עם המאורע B בפתרון הרשמי של "לתפוס את הרכבת", בתוספת אינדוקציה) שלכל k מתקיים $\Pr[A_k] = (\Pr[A])^k$. האפשרות השניה היא לשים לב שמתקיים $\Pr[\neg A] \geq \Pr[\neg A \wedge (X_k = -k)] = \Pr[\neg A|X_k = -k]\Pr[X_k = -k]$. מכאן משתמשים בתובנה המרכזית של הפתרון של "לתפוס את הרכבת" ומסיקים שמתקיים $\Pr[\neg A|X_k = -k] = \Pr[\neg A_k]$, בעוד שמתקיים $\Pr[X_k = -k] = (\frac{2}{3})^k > 0$.

עתה נכתוב הוכחה, שעבור k גדול מספיק (שנבחר בסוף) תתן לנו שאכן מתקיים $\Pr[A_k] < 1$. נגדיר מאורעות C_i עבור $i \in \mathbb{N}$, כאשר C_i הוא המאורע שמתקיים $X_i \geq k$, שהוא שקול למאורע $\sum_{j=1}^{i} Z_j \geq \frac{k}{2} + \frac{i}{2}$. בפרט מתקיים $A_k = \bigvee_{i\in\mathbb{N}} C_i$ ולכן $\Pr[A_k] \leq \sum_{i\in\mathbb{N}} \Pr[C_i]$. עתה נותר לחסום את המחוברים בצד ימין.

עבור $i < k$, מתקיים בפשטות $\Pr[C_i] = 0$. עבור $i \geq k$, לפי חסימת סטיות גדולות (החסם השני מהשיעור) מתקיים $\Pr[\sum_{j=1}^{i} Z_j \geq \frac{k}{2} + \frac{i}{2}] \leq \Pr[\sum_{j=1}^{i} Z_j \geq \frac{i}{2}] \leq e^{-i/18}$. כל שנותר הוא לבחור k שעבורו מתקיים $\sum_{i=k}^{\infty} e^{-i/18} < 1$. קיים לנו k כזה כי הטור $\sum_{i=0}^{\infty} e^{-i/18}$ הוא טור חזקות מתכנס. עבור ה־k שבחרנו נקבל $\Pr[A_k] \leq \sum_{i\in\mathbb{N}} \Pr[C_i] \leq \sum_{i=k}^{\infty} e^{-i/18} < 1$, כנדרש.

הערה: פתרון אלגנטי אחר לשאלה משתמש בחוק המספרים הגדולים החזק. אם מתקיים $\Pr[A] = 1$ אז (לפי הטיעונים שצוינו למעלה) יתקיים לכל k גם $\Pr[A_k] = 1$, ולכן בהסתברות 1 הסדרה תקיים $\overline{\lim}_{l\to\infty} X_l = +\infty$ (שימו לב שזה גבול עליון, כמובן שלא מובטח שהסדרה מתכנסת). אבל מסקנה של חוק המספרים הגדולים החזק (בגלל ש־X_i ניתנים להצגה של סכום מ"מ ב"ת בעלי התפלגויות זהות ותוחלת שלילית) היא שבהכרח בהסתברות 1 יתקיים $\lim_{k\to\infty} X_l = -\infty$ (הפעם גבול אמיתי), שזו סתירה.

תתי־קבוצה מקריים

נסמן את תתי־הקבוצה המקריים ב־$A_1, \ldots, A_{2n}$, ולכל $1 \leq i \leq 2n$ נסמן ב־X_i את משתנה האינדיקטור עבור המאורע ש־A_i מכילה לפחות איבר אחד מ־$\{1, \ldots, n\}$ שמוכל בלפחות 40 מהקבוצות $A_1, \ldots, A_{i-1}$. אם בסוף נבחר את תתי־הקבוצה $\{A_i|X_i = 0\}$, הרי שאלו יקיימו את התנאי הנדרש על מספר המופעים של איברי $\{1, \ldots, n\}$, ולכן עלינו להוכיח שבסיכוי חסום ע"י $e^{-\Theta(n)}$ בלבד יהיו פחות מ־n קבוצות כאלו, דבר השקול לביטוי $X = \sum_{i=1}^{2n} X_i > n$.

לפי ספירה פשוטה, מספר האיברים מ־$\{1, \ldots, n\}$ המשתתפים בלפחות 40 קבוצות מ־$A_1, \ldots, A_{i-1}$ לעולם אינו עולה על $\frac{3}{20}n$. על כן, אפילו ש־X_i תלוי ב־$X_1, \ldots, X_{i-1}$, יתקיים $\Pr[X_i = 1|X_1, \ldots, X_{i-1}] \leq \frac{9}{20}$ (לכל סדרת ערכים אפשרית עבור $X_1, \ldots, X_{i-1}$).

על מנת לחסום את $X=\sum_{i=1}^{2n}X_i$ נגדיר $Y_1,\ldots,Y_{2n}$ אשר יהיו ב"ת זה בזה לחלוטין (אבל תלויים ב־$X_1,\ldots,X_{2n}$) ושעבורם יתקיים $Y_i\geq X_i$ וכן $\Pr[Y_i]=\frac{9}{20}$. נבנה את אלו באינדוקציה, את Y_i נגדיר לאחר שהוגרלו $X_1,\ldots,X_i$ ו־$Y_1,\ldots,Y_{i-1}$ (ונדאג לאי תלות של ההסתברויות עבור Y_i בערכים של $Y_1,\ldots,Y_{i-1}$).

בהינתן הערכים $X_1=\alpha_1,\ldots,X_{i-1}=\alpha_{i-1}$ ו־$p_i=\Pr[X_i=1|X_1=\alpha_1,\ldots,X_{i-1}=\alpha_{i-1}]$, נבדוק עתה את ערך X_i. אם $X_i=1$ אז נגדיר $Y_i=1$, ואם $X_i=0$ אז בהסתברות $\frac{11}{20(1-p_i)}$ נגדיר $Y_i=0$ ובהסתברות $\frac{9-20p_i}{20(1-p_i)}$ נגדיר $Y_i=1$. הדבר לשים לב אליו הוא ש־Y_i יהיה שווה ל־1 בהסתברות $\frac{9}{20}$ באופן ב"ת בערכי $X_1,\ldots,X_{i-1}$ או ערכי $Y_1,\ldots,Y_{i-1}$. לכן קבוצת כל ה־Y_i היא ב"ת (במובן שמתקיים למשל $\Pr\left[\bigwedge_{j=1}^{i}(Y_j=1)\right]=\left(\frac{9}{20}\right)^i$).

עתה ניתן לחסום את $\Pr[X>n]$ ע"י $\Pr[\sum_{i=1}^{2n}Y_i>n]$, וזה חסום ע"י $e^{-\Theta(n)}$ באמצעות חסימת סטיות גדולות.

שוויון הנפש

נבדוק תהליך שבו מגרילים קבוצה B באופן יוניפורמי מבין 2^n תתי־הקבוצה האפשריים, וננתח את הערכים האפשריים עבור המ"מ $X_1,\ldots,X_k$, כאשר מגדירים $X_i=|A_i\cap B|$. בסופו של דבר הרעיון יהיה להשתמש בטיעון "שובך יונים" עבור סדרות הערכים האפשריות של המ"מ הנ"ל, אבל אי אפשר לעשות את זה עבור X_i שבאופן תאורטי יכולים לקבל כל ערך בין 0 ל־n.

ננתח עבור $1\leq i\leq k$ קבוע את התפלגות X_i: תוחלת המשתנה היא $|A_i|/2$. מחסימת סטיות גדולות, מתקיים $\Pr[X_i>|A_i|/2+\sqrt{|A_i|\log(n)}]<e^{-2|A_i|\log(n)/|A_i|}<\frac{1}{4k}$ וכן $\Pr[X_i<|A_i|/2-\sqrt{|A_i|\log(n)}]<\frac{1}{4k}$ (עבור ההשוואה ל־$\frac{1}{4k}$ אנחנו משתמשים בכך שבפרט מתקיים $k<2n$, ומניחים ש־N_α נקבע להיות גדול דיו שיתקיים $e^{-2\log(n)}<n^{-2}<\frac{1}{8n}$). על כן, בהסתברות לפחות $\frac{1}{2}$, אף אחד מהמאורעות המתוארים למעלה לא מתקיים לאף $1\leq i\leq k$. זה אומר שלפחות $2^n/2$ מהבחירות האפשריות עבור B הן כאלה שעבורן לכל X_i יש לכל היותר $1+2\sqrt{|A_i|\log(n)}\leq 4\sqrt{n\log(n)}$ ערכים (שוב בהנחה ש־$n>N_\alpha$ הוא גדול דיו).

עכשיו נשתמש בעקרון שובך היונים עבור $2^n/2$ הבחירות הנ"ל. מספר סדרות הערכים האפשריות עבור $X_1,\ldots,X_k$ חסום ע"י $(4\sqrt{n\log(n)})^k=2^{k(\log(n)/2+\log\log(n)/2+2)}\leq 2^{(\alpha/2+o(1))n}$. מכיוון שנתון ש־$\alpha<2$, עבור n גדול דיו (כתוצאה מבחירה מתאימה של N_α) הביטוי הנ"ל יהיה קטן מ־$2^n/2$, ולכן יהיו לפחות שתי בחירות של B עם אותה סדרת ערכים עבור $|A_1\cap B|,\ldots,|A_k\cap B|$, כנדרש.

קליקים בממוצע

נראה כאן באינדוקציה שלכל k ו־$0<c<1$ קיים $\alpha(c,k)>0$ כך שעבור n גדול דיו, בהסתברות של לכל היותר $e^{-\alpha n}$, מספר ה־k־קליקים יהיה פחות מ־$(1-c)2^{-\binom{k}{2}}\binom{n}{k}$ או יותר מ־$(1+c)2^{-\binom{k}{2}}\binom{n}{k}$. טענת השאלה נובעת מהצבת $c=\frac{1}{2}$. בסיס האינדוציה, $k=1$, הוא טריביאלי (תמיד יהיו בדיוק n "1־קליקים").

נניח שהטענה הוכחה עבור $k-1$. ראשית נראה שלכל $0<c<1$ ו־n גדול דיו, בהסתברות של לכל היותר $e^{-\beta n}$, עבור $\beta(c,k)>0$ מתאים, ישנה קבוצה "רעה" U בת $k-1$ צמתים כך שמספר הצמתים $v\notin U$ אשר מחוברים לכל איברי U לא נמצא בין $(1-c)2^{1-k}(n+1-k)$ ל־$(1+c)2^{1-k}(n+1-k)$ (כלומר חורג מהאמור בטענה). נסמן ב־$A_{v,U}$ את המאורע ש־$v\notin U$ מחובר לכל צמתי U. $\Pr[A_{v,U}]=2^{1-k}$, ועבור U קבוע, המאורעות $\{A_v:v\in V\setminus U\}$ הם ב"ת לחלוטין זה בזה. לכן, מחסימת סטיות גדולות, קיים $\beta'(c,k)>0$ כך שבהסתברות לכל היותר $e^{-\beta' n}$ מספר הצמתים $v\notin U$ המחוברים לכל איברי U לא נמצא בין $(1-c)2^{1-k}(n+1-k)$ ל־$(1+c)2^{1-k}(n+1-k)$. מכיוון שמספר הקבוצות U הרלוונטיות הוא $\binom{n}{k-1}$, ניתן להפעיל את חסם האיחוד ולקבל כי קיים $\beta(c,k)>0$ כך שעבור n גדול דיו בהסתברות של לכל היותר $e^{-\beta n}$ קיימת קבוצה "רעה".

אם עתה נספור את כל הזוגות של $(k-1)$־קליק פלוס צומת נוסף המשלים אותו ל־k־קליק, נקבל מהטענה לעיל והנחת האינדוקציה (מופעלים שניהם עם $c/3$) שעבור n גדול דיו ההסתברות שמספר הזוגות לא יהיה בין $(1-c)2^{-\binom{k-1}{2}}2^{1-k}\binom{n}{k-1}(n+1-k)=(1-c)2^{-\binom{k}{2}}k\binom{n}{k}$ לבין $(1+c)2^{-\binom{k}{2}}k\binom{n}{k}$ היא חסומה ע"י

$e^{-\alpha(c/3,k-1)n} + e^{-\beta(c/3,k)n}$ (שכן $(1+c/3)^2 \leq 1+c$ ו־$(1-c/3)^2 \geq 1-c$ תחת ההנחה ש־$0 < c < 1$).
אם נשים לב שמספר הזוגות הנ"ל הוא בדיוק k פעמים מספר ה־k־קליקים בגרף, נוכל לסיים ע"י ההצבה $\alpha(c,k) = \frac{1}{2}\min\{\alpha(c/3,k-1), \beta(c/3,k)\}$.

מרטינגלים

מהמר עם זמן

הראו שלא קיים מרטינגל בעל אורך לא סופי $X_0, X_1, \ldots$ שמקיים (בהסתברות 1) את $X_0 = 0$, לכל i מקיים את $X_i \geq -C$ עבור קבוע גלובלי C כל שהוא, ובנוסף מקיים $\lim_{i\to\infty} \Pr[X_i \geq 1] = 1$ (בהסתברויות לא מותנות). אפשר לחשוב על זה כעל גרסה של "מרטינגל המהמר החכם" מההרצאה, רק שכאן יש למהמר זמן אינסופי ורק התקציב שלו מוגבל (מסתבר שגם ככה אי אפשר להרוויח).

הילוך מקרי על הקוביה

על הקוביה הבוליאנית $\{0,1\}^n$ נגדיר הילוך מקרי מהראשית: $\underline{x}^{(0)}$ יהיה הווקטור שכולו אפסים, ובהינתן $\underline{x}^{(i)}$, נגדיר את $\underline{x}^{(i+1)}$ ע"י זה שנבחר באופן יוניפורמי (וב"ת בבחירות קודמות) אינדקס $1 \leq j_i \leq n$, ואז נהפוך את ערכו של האיבר ה־j_i ב־$\underline{x}^{(i)}$. נסמן ב־d_i את המרחק מהראשית של $\underline{x}^{(i)}$, ז"א את מספר האחדות שבו. הראו שמתקיים:

$$\Pr\Big[d_n < \frac{1}{2}\mathrm{E}[d_n]\Big] \leq 2^{-\Omega(n)}$$

הערה: אפשר לפתור זאת מבלי לחשב במדויק את $\mathrm{E}[d_n]$, אבל כמובן שצריך לדאוג לאיזה שהוא חסם על גודל התוחלת הנ"ל.

מספרים מכוסים

הראו שלכל k קיים α_k עם התכונה הבאה: כאשר בוחרים באופן יוניפורמי וב"ת n מספרים מ־$\{1,\ldots,n\}$, הסיכוי שאין אף מספר שנבחר לפחות k פעמים חסום ע"י $2^{-\alpha_k n}$.

מרחק מקבוצת נקודות

נניח ש־$A \subseteq \{0,1\}^n$ היא קבוצה בת לפחות $\frac{1}{100}2^n$ נקודות. הראו שלפחות $\frac{99}{100}2^n$ מנקודות הקוביה $\{0,1\}^n$ נמצאות במרחק שאינו עולה על $8\sqrt{n}$ מ־A, כאשר המרחק בין שתי נקודות מוגדר ע"י מספר הקורדינטות שבהן הן נבדלות (מרחק Hamming ללא נרמול).

לסחוט את הלימון

נתון ש־$X_0, \ldots, X_m$ הוא מרטינגל מעל מרחב הסתברות בדיד (כך שלא צריך להסתבך עם הגדרות של מרחבים רציפים) שעבורו $X_0 = 0$ (בהסתברות 1). כמו כן נתון שלכל $1 \leq i \leq m$ וסדרת ערכים $\alpha_0, \ldots, \alpha_{i-1}$ שמתקיים עבורם $\Pr[X_0 = \alpha_0, \ldots, X_{i-1} = \alpha_{i-1}] > 0$, קיים $p = p_{\alpha_0,\ldots,\alpha_{i-1}} \in [0,1]$ שעבורו מתקיים $\Pr[X_{i-1} - p \leq X_i \leq X_{i-1} + 1 - p | X_0 = \alpha_0, \ldots, X_{i-1} = \alpha_{i-1}] = 1$. במילים אחרות, לכל "הסטוריה אפשרית" של המרטינגל לפני X_i, תחום הערכים האפשריים של X_i מוכל בקטע באורך 1 (שכמובן מכיל את הערך של X_{i-1}, כי אחרת זה לא יכול להיות מרטינגל). הראו שלכל $a > 0$ מתקיים $\Pr[X_m > a] < e^{-2a^2/m}$.

הערה: הגרסה הזו מקבילה למשפט הריכוז על סכום משתנים ב"ת מוטים שנלמד בהרצאה (המשפט השני בפרק המתאים). גרסה זו מחזירה את המקדם בתוך החזקה ל"מקומו הראוי" כפי שמופיע באותו משפט ריכוז.

לימונדה

נניח ש־$X_0,\ldots,X_m$ הוא מרטינגל חשיפה של פונקציה f שמוגדר לפי התפלגות μ מעל מרחב הפונקציות מ־$\mathcal{D}$ ל־$\mathcal{R}$, עם סדר חשיפה $\emptyset = \mathcal{D}_0 \subseteq \cdots \subseteq \mathcal{D}_m = \mathcal{D}$. נניח גם ש־$\mu$ היא התפלגות מכפלה (לכל $d \in \mathcal{D}$ מגרילים את $C(d) \in \mathcal{R}$ באופן ב"ת בהגרלות האחרות), וגם שלכל $1 \le i \le m$ ולכל $C_1, C_2 : \mathcal{D} \to \mathcal{R}$ שנבדלות ביניהם רק על $\mathcal{D}_i \setminus \mathcal{D}_{i-1}$ מתקיים $|f(C_1) - f(C_2)| \le 1$. הראו שלכל $1 \le i \le n$ ולכל $\tilde{C} : \mathcal{D} \to \mathcal{R}$ המתקבל בהסתברות חיובית מתקיים $\max\{X_i(\hat{C}) : \hat{C}|_{\mathcal{D}_{i-1}} = \tilde{C}|_{\mathcal{D}_{i-1}}\} \le 1 + \min\{X_i(\hat{C}) : \hat{C}|_{\mathcal{D}_{i-1}} = \tilde{C}|_{\mathcal{D}_{i-1}}\}$.

הערה: מהתוצאה הזו אפשר להוכיח, בשיטה מאוד דומה לשאלה "לסחוט את הלימון" (תוך ניתוח המרטינגל המוגדר ע"י $X'_i = X_i - X_0$), שבתנאים אלו מתקיים חסם ההסתברות המשופר $\Pr[X_m > X_0 + a] < e^{-2a^2/m}$.

בחירה של תתי־קבוצה

נתון ש־A היא ת"ק מקרית של $\{1,\ldots,n\}$. לא נתון כלום על מרחב ההסתברות שלפיו בוחרים את A, פרט לכך שזוהי בהסתברות 1 קבוצה בת k איברים שונים זה מזה (k הוא קבוע נתון), וכן שלכל $1 \le i \le n$ מתקיים $\Pr[i \in A] = \frac{k}{n}$. הראו שעבור $(1-o(1))2^n$ מתתי הקבוצות $S \subseteq \{1,\ldots,n\}$ מתקיים $\Pr[A \subseteq S] = 2^{-k} \pm o(1)$.

רמז: ניתן להסתכל על $\Pr[A \subseteq S]$ כעל כמות התלויה ב־S, ולנתח את ההתפלגות עבור בחירה מקרית יוניפורמית של S.

אי אפשר להתחמק

מגרילים קבוצה $U \subseteq \{1,\ldots,n\}$ ע"י כך שכל $1 \le i \le n$ נבחר להיות ב־U בהסתברות p (עבור קבוע $0 < p < 1$) באופן ב"ת בכל הבחירות האחרות. כמו כן נתון שיש אלגוריתם $\mathcal{A}$ שבוחר קבוצה $A \subseteq \{1,\ldots,n\}$ באופן "אדפטיבי": בשלב ה־i האלגוריתם בוחר האם $i \in A$, והבחירה הזו יכולה להיות תלויה בזהות של $U \cap \{1,\ldots,i-1\}$. הראו שלכל $a > 0$ מתקיים $\Pr[|A \cap U| - p|A| > a] < e^{-a^2/2n}$, בלי קשר לאופן שבו האלגוריתם $\mathcal{A}$ מבצע את הבחירות שלו.

הערות: הגודל $|A|$ עצמו יכול להתפלג באופנים שלא מקיימים משפטי ריכוז. משפט הריכוז שאתם צריכים להוכיח הוא על הגודל של $|A \cap U|$ ביחס ל־$|A|$. מותר להניח ש־$\mathcal{A}$ הוא דטרמיניסטי. המשמעות היא שאפשר להסתכל על $\mathcal{A}$ כעל סדרה של פונקציות $\mathcal{A}_i : \mathcal{P}(\{1,\ldots,i-1\}) \to \{0,1\}$ עבור $1 \le i \le n$, ובסוף מתקיים $A = \{i : \mathcal{A}_i(U \cap \{1,\ldots,i-1\}) = 1\}$.

צביעת גרף מושרה על קבוצה מקרית

נניח ש־$G = (V,E)$ הוא גרף שמספר הצביעה שלו הוא בדיוק 1000. נניח ש־V' היא תת קבוצה של V שנבחרת אקראית ויוניפורמית (כל צומת ב־V נבחר עבור V' בהסתברות $\frac{1}{2}$ באופן ב"ת). יהי G' תת הגרף המושרה על V'. הראו שבסיכוי $\frac{99}{100}$ לפחות מתקיים $\chi(G') \ge 400$. כדאי קודם להראות שמתקיים $\mathrm{E}[\chi(G')] \ge 500$.

פתרונות לתרגילים על מרטינגלים

מהמר עם זמן

נניח שקיים מרטינגל $X_0, X_1, \ldots$ כזה עבור קבוע C מתאים. לפי תנאי השאלה, בפרט קיים N כך שלכל $n \ge N$ מתקיים $\Pr[X_n \ge 1] > C/(C+1)$. אבל אז בפרט $\Pr[X_n < 0] < 1/(C+1)$, ומתקיים כזכור $X_n \ge -C$ תמיד. סכימה אם כך תתן לנו $\mathrm{E}[X_n] > 0$, אולם זוהי סתירה, מכיוון שבמרטינגל מתקיים $\mathrm{E}[X_n] = \mathrm{E}[X_0]$, וכזכור הנחנו שמתקיים $X_0 = 0$ תמיד.

הערה: עם קצת יותר עבודה אפשר גם להראות שמתקיים $\lim_{n\to\infty} \Pr[X_n \ge \alpha] = 0$ לכל $\alpha > 0$. זיכרו שזה נכון לכל מרטינגל, כולל אלו עם "תנאי עצירה" (שפשוט נהיים קבועים אחרי שהתנאי מתקיים).

הילוך מקרי על הקוביה

נגדיר מרטינגל חשיפה $D_0,\dots,D_n$ עבור המרחק d_n, כאשר בצעד ה־i חושפים את j_i (ולכן את $\underline{x}^{(0)},\dots,\underline{x}^{(i)}$). שימו לב שבד"כ D_i אינו שווה ל־d_i עבור $i<n$ (אבל כמובן $D_n=d_n$). לא קשה לראות שהמרטינגל מקיים את תנאי ליפשיץ עם קבוע ליפשיץ 2, ולכן ממשפט אזומה מתקיים $\Pr[D_n<\mathrm{E}[D_n]-\alpha n]\le 2^{-\Omega(n)}$ לכל $\alpha>0$. על מנת להשלים את ההוכחה על כן צריך רק להראות שמתקיים $\mathrm{E}[D_n]=\Omega(n)$.

הסיכוי שאיבר i נבחר בדיוק פעם אחת הוא לפחות $1-\left(1-\frac{1}{n}\right)^n-\binom{n}{2}\cdot\frac{1}{n^2}$ (הסיכוי שיבחר לפחות פעם אחת הוא $1-\left(1-\frac{1}{n}\right)^n$ והסיכוי שיבחר פעמיים ומעלה הוא לכל היותר $\binom{n}{2}\cdot\frac{1}{n^2}$ מחסם האיחוד), עבור n גדול דיו ערך זה הוא לפחות $\frac{1}{10}$, ולכן מלינאריות התוחלת עבור n גדול דיו מתקיים $\mathrm{E}\left[D_n\right]\ge\frac{n}{10}$.

מספרים מכוסים

נניח ש־$C:\{1,\dots,n\}\to\{1,\dots,n\}$ תסמן את סדרת n המספרים שבחרנו, ונסמן ב־$f(C)$ את הפונקציה שאומרת כמה מספרים נבחרו לפחות k פעמים. ראשית נחסום את התוחלת שלה: לכל מספר $1\le i\le n$, הסיכוי ש־i נבחר לפחות k פעמים הוא לפחות הסיכוי ש־i נבחר בדיוק k פעמים, $\binom{n}{k}n^{-k}(1-\frac{1}{n})^{n-k}$. מה שדרוש לנו זה שעבור $n>k$ הביטוי חסום מלמטה ע"י קבוע מתאים β_k (עבור k קבוע הביטוי שואף ל־$1/ek!$). אפשר לראות את $f(C)$ כסכום של n משתני האינדיקטור עבור כל $1\le i\le n$, ולכן מלינאריות התוחלת $\mathrm{E}_C[f(C)]\ge\beta_k n$.

עתה נבנה מרטינגל חשיפה של $f(C)$ כאשר חושפים את C איבר־איבר, ז"א $\mathcal{D}_i=\{1,\dots,i\}$. לא קשה לראות שהפונקציה הזו היא ליפשיץ ביחס לחשיפה, ז"א ששינוי של C במקום אחד משנה את כמות המספרים שמופיעים לפחות k פעמים בלא יותר מאחד (השינוי יכול רק להוסיף או להוציא מספר בודד מקבוצת המספרים הזו). מכאן שגם המרטינגל $X_0,\dots,X_m$ מקיים $|X_i-X_{i-1}|\le 1$, ולכן לפי אי שוויון אזומה מתקיים החסם $\Pr[f(C)=0]\le\Pr[f(C)\le\mathrm{E}[f(C)]-\beta_k n]\le e^{-\beta_k^2 n/2}$.

מרחק מקבוצת נקודות

הרעיון כאן הוא להראות שכאשר מגרילים באופן יוניפורמי נקודה $x=(x_1,\dots,x_n)$ מ־$\{0,1\}^n$, בהסתברות לפחות $\frac{99}{100}$ המרחק שלה מ־A לא יעלה על $8\sqrt{n}$. לשם כך נסתכל על הנקודה המקרית כעל פונקציה מקרית באופן יוניפורמי מ־$\{1,\dots,n\}$ ל־$\{0,1\}$, ונבנה מרטינגל חשיפה ביחס לפונקצית המרחק של x מ־A. המרטינגל "יחשוף" קורדינטה אחת בכל שלב, ז"א שהחשיפה תיעשה ביחס לתחום $\mathcal{D}_i=\{1,\dots,i\}$ לכל $0\le i\le n$. שימו לב שבמרטינגל $X_0,\dots,X_n$ המתקבל כך, הערך של X_i אינו מקבל את המרחק של הצמצום $(x_1,\dots,x_i)$ מהצמצום המתאים של נקודות A. הערך של X_i שווה לתוחלת המותנה של המרחק הכולל בהינתן הערכים של $x_1,\dots,x_i$, בהתאם להגדרה של מרטינגל חשיפה.

בחירת ערכי x על הקורדינטות היא ב"ת, ולכן לא קשה לראות שמתקיים תנאי ליפשיץ עבור המרטינגל מקיום תנאי ליפשיץ עבור התחומים: אם שתי נקודות נבדלות ביניהן רק על $\mathcal{D}_i\setminus\mathcal{D}_{i-1}=\{i\}$, אז המרחקים שלהן מ־$A$ בוודאי לא נבדלים ביותר מ־1. מכאן נובע ע"י אי שוויון Azuma שהסיכוי שמרחק זה יהיה קטן מהתוחלת שלו ביותר מ־$4\sqrt{n}$ אינו עולה על $e^{-8}<\frac{1}{100}$. מצד שני, בסיכוי לפחות $\frac{1}{100}$ המרחק המתקבל הוא 0, כי זהו הסיכוי ש־$x\in A$, ולכן תוחלת המרחק של x מ־A אינה עולה על $4\sqrt{n}$. מכאן נובע שבסיכוי לפחות $\frac{99}{100}$ (ע"י שימוש נוסף באי שוויון Azuma), המרחק של x מ־A אינו עולה בעצמו על $8\sqrt{n}$.

לסחוט את הלימון

ראשית נתייחס להוכחה של משפט הריכוז על משתנים מוטים מחוברת ההרצאות. נוכיח שלכל $\lambda>0$ מתקיים $\mathrm{E}[e^{\lambda X_m}]<e^{\lambda^2 m/8}$. אחרי שמוכיחים זאת מציבים $\lambda=\frac{4a}{m}$ ומשתמשים באי־שוויון מרקוב, כפי שנעשה שם.

הוכחת הטענה משתמשת בשיטה דומה לחסם על תוחלת $e^{\alpha X_m}$ בהוכחה של משפט אזומה. ראשית נוכיח שלכל מ"מ Z שמקיים $\mathrm{E}[Z]=0$ ו־$\Pr[-p\le Z\le 1-p]=1$ עבור $p\in[0,1]$ מתאים, מתקיים $\mathrm{E}[e^{\lambda Z}]<e^{\lambda^2/8}$.

לשם כך נשתמש בקמירות של הפונקציה $f(z)=e^{\lambda z}$ על התחום $[-p,1-p]$. כזכור מתקיים בתחום הזה $f(z)\le g(z)$, כאשר $g(z)$ היא הפונקציה הלינארית העוברת דרך $(-p,f(-p))$ ו־$(1-p,f(1-p))$. אם נרשום את הפונקציות במפורש, נקבל שעבור כל $-p\le z\le 1-p$ מתקיים

$$e^{\lambda z}\le\left(pe^{\lambda(1-p)}+(1-p)e^{-\lambda p}\right)+\left(e^{\lambda(1-p)}-e^{-\lambda p}\right)z$$

ע"י שימוש בלינאריות התוחלת, נובע מזה $\mathrm{E}[f(Z)]\le\mathrm{E}[g(Z)]=pe^{\lambda(1-p)}+(1-p)e^{-\lambda p}\le e^{\lambda^2/8}$, כאשר אי־השוויון מימין מופיע (אומנם בלי הוכחה) בהוכחת משפט הריכוז המקורית. עתה אנחנו יכולים להוכיח באינדוקציה את הטענה הדרושה על $\mathrm{E}[e^{\lambda X_m}]$: נגדיר $Y_i=X_i-X_{i-1}$ לכל $1\le i\le m$, ונכתוב

$$\mathrm{E}\left[e^{\lambda X_m}\right]=\mathrm{E}\left[e^{\lambda Y_m}e^{\lambda X_{m-1}}\right]\le e^{\lambda^2/8}\mathrm{E}\left[e^{\lambda X_{m-1}}\right]\le\cdots\le e^{\lambda^2 m/8}$$

זוהי בדיוק האינדוקציה שנעשתה בחסימת התוחלת בהוכחה של משפט אזומה, וגם כאן משתמשים בכך שתכונות Y_m, ובפרט גם החסם על התוחלת של $e^{\lambda Y_m}$, מתקיימים לכל התניה מהצורה $X_0=\alpha_0,\ldots,X_{i-1}=\alpha_{i-1}$.

לימונדה

על מנת להראות את המבוקש ניזכר איך בחוברת ההרצאות, בהוכחה של השיטה הנפוצה לחסימת מרטינגל חשיפה, מגדירים בהינתן $\tilde{C}:\mathcal{D}\to\mathcal{R}$ מרחב הסתברות על זוגות $C_1,C_2:\mathcal{D}\to\mathcal{R}$. כאן, בהינתן $\tilde{C}$ ו־$\hat{C}$ שעבורם $\hat{C}|_{\mathcal{D}_{i-1}}=\tilde{C}|_{\mathcal{D}_{i-1}}$, נגדיר מרחב הסתברות מעל שלישיות $C_0,C_{\min},C_{\max}$ באופן הבא: נגריל את C לפי μ (כמו בהוכחה שם), וראשית נקבע את C_0 לפי $C_0(d)=\hat{C}(d)$ עבור $d\in\mathcal{D}_i$ ו־$C_0(d)=C(d)$ עבור $d\in\mathcal{D}\setminus\mathcal{D}_i$. נשים לב שבדומה להוכחה בחוברת ההרצאה, מתקיים $\mathrm{E}[f(C_0)]=X_i(\hat{C})$.

עבור הגדרת $C_{\max}$ ו־$C_{\min}$, נגדיר את $S_{C,\tilde{C}}$ להיות קבוצת המבנים $C':\mathcal{D}\to\mathcal{R}$ שעבורם $C'|_{\mathcal{D}_{i-1}}=\tilde{C}|_{\mathcal{D}_{i-1}}$ וכן $C'|_{\mathcal{D}\setminus\mathcal{D}_i}=C|_{\mathcal{D}\setminus\mathcal{D}_i}$. נשים לב שבפרט $C_0\in S_{C,\tilde{C}}$. עתה נגדיר את $C_{\max}$ להיות איבר מתוך $S_{C,\tilde{C}}$ שעבורו $f(C_{\max})=\max\{f(C'):C'\in S_{C,\tilde{C}}\}$ (אם יש יותר מאיבר אחד שמקבל את המקסימום אז בוחרים אחד מהם שרירותית), ונגדיר את $C_{\min}$ להיות איבר מ־$S_{C,\tilde{C}}$ שעבורו $f(C_{\min})=\min\{f(C'):C'\in S_{C,\tilde{C}}\}$.

לפי הנתונים של השאלה, תמיד מתקיים $f(C_{\max})\le 1+f(C_{\min})$. זאת כי הדבר מתקיים לכל זוג איברים מ־$S_{C,\tilde{C}}$, שכן כל איברי הקבוצה הזו נבדלים על $\mathcal{D}_i\setminus\mathcal{D}_{i-1}$ בלבד. לכן, כאשר מגרילים את C, מתקיים $\mathrm{E}[f(C_{\max})]\le 1+\mathrm{E}[f(C_{\min})]$. לבסוף, בגלל שלכל C מתקיים $f(C_{\min})\le f(C_0)\le f(C_{\max})$, לכל $\hat{C}$ שעבורו $\hat{C}|_{\mathcal{D}_{i-1}}=\tilde{C}|_{\mathcal{D}_{i-1}}$ יתקיים גם $\mathrm{E}[f(C_{\min})]\le\mathrm{E}[f(C_0)]=X_i(\hat{C})\le\mathrm{E}[f(C_{\max})]$. מכך נובע המבוקש:

$$\begin{aligned}\max\{X_i(\hat{C}):\hat{C}|_{\mathcal{D}_{i-1}}=\tilde{C}|_{\mathcal{D}_{i-1}}\} &\le \mathrm{E}[f(C_{\max})]\\ &\le 1+\mathrm{E}[f(C_{\min})]\le 1+\min\{X_i(\hat{C}):\hat{C}|_{\mathcal{D}_{i-1}}=\tilde{C}|_{\mathcal{D}_{i-1}}\}\end{aligned}$$

בחירה של תתי־קבוצה

לכל $S\subseteq\{1,\ldots,n\}$ נגדיר $p(S)=\Pr[A\subseteq S]$, ונגדיר את $X_0,\ldots,X_n$ להיות מרטינגל החשיפה של S כאשר מחשיבים אותה כפונקציה מקרית מ־$\{1,\ldots,n\}$ ל־$[0,1]$. באופן יותר פורמלי, עבור $T\subseteq\{1,\ldots,n\}$, המ"מ X_i יוגדר לפי התוחלת המותנה $X_i(T)=\mathrm{E}_{S\subseteq\{1,\ldots,n\}}[p(S)|T\cap\{1,\ldots,i\}=S\cap\{1,\ldots,i\}]$.

ברור שמתקיים $X_0=\mathrm{E}[p(S)]=2^{-k}$, כי הסיכוי עבור $A\subseteq S$ הוא 2^{-k} לכל A בגודל k. כמו כן מתקיים בהסתברות 1 התנאי $|X_i-X_{i-1}|\le\frac{k}{n}$: זה נובע מ"תנאי ליפשיץ" מתאים על הפונקציה $p(s)$, שהרי אם S

ו־T נבדלות ביניהן רק על הקואורדינטה ה־i, נניח בלי הגבלת הכלליות ש־$T = S \setminus \{i\}$. במקרה זה מתקיים $0 \le p(S) - p(T) \le \Pr_A[i \in A] = \frac{k}{n}$ לפי הנתון על המרחב ההסתברות שלפיו בוחרים את A. מאי שוויון Azuma נובע עתה שבהסתברות $1 - o(1)$ (עבור n גדול דיו ביחס ל־k) המרחק בין X_s ל־X_0 הוא $o(1)$ כנדרש.

אי אפשר להתחמק

על מנת "לחסוך מקום", נסמן $U_i = U \cap \{1, \ldots, i\}$ עבור $0 \le i \le n$, ובפרט $U_0 = \emptyset$ ו־$U_n = U$. נגדיר עכשיו את המ"מ $X_0, \ldots, X_n$ מעל מרחב ההסתברות של בחירת הקבוצה U (באופן פורמלי אלו פונקציות ממשיות $X_i : \mathcal{P}(\{1, \ldots, n\}) \to \mathbb{R}$), לפי $X_i(U) = |A \cap U_i| - p|A \cap \{1, \ldots, i\}|$. ברור שמתקיים $X_0 = 0$ בהסתברות 1, וכן מתקיים $X_n = |A \cap U| - p|A|$. עתה נוכיח שהסדרה הנ"ל היא מרטינגל, ושמתקיים $|X_i(U) - X_{i-1}(U)| \le 1$ לכל U ו־$1 \le i \le n$.

נשים לב ש־$X_0(U), \ldots, X_{i-1}(U)$ תלויים אך ורק ב־U_{i-1}, ישירות מההגדרות של המ"מ יחד עם זה שהחיתוך $A \cap \{1, \ldots, i-1\}$ תלוי אך ורק ב־$U \cap \{1, \ldots, i-2\}$ (שימו לב להגדרות של $\mathcal{A}_1, \ldots, \mathcal{A}_{i-1}$ בהערות לתרגיל). נראה שלכל ערך אפשרי V של U_{i-1} מתקיים $\mathrm{E}[X_i(U)|U_{i-1} = V] = X_{i-1}(V)$, כאשר $X_{i-1}(V)$ מסמן את הערך המשותף של $X_{i-1}(U)$ לכל ה־U המקיימים $U_{i-1} = V$. זה מוכיח שהמדובר במרטינגל, כי התוחלת המותנה על סדרת הערכים של $X_0, \ldots, X_{i-1}$ תהיה ממוצע משוקלל לפי כל ה־V שנותנים את הסדרה הזו. את השלמת ההוכחה מבצעים לפי פיצול למקרים: אם $\mathcal{A}_i(V) = 0$, אז $X_i = X_{i-1}$ (כי A לא יכיל את i). אם $\mathcal{A}_i(V) = 1$, אז בהסתברות p יתקיים $X_i = X_{i-1} + 1 - p$ (אם i נבחר להיות ב־U), ובהסתברות $1 - p$ יתקיים $X_i = X_{i-1} - p$ (אם i לא נבחר להיות ב־U). שני אלו שוב נותנים תוחלת מותנה ששווה ל־X_{i-1}.

לחסימת $|X_{i-1}(U) - X_i(U)|$, פשוט מסתכלים על כל האפשרויות שנזכרו כבר בפסקה הקודמת: בפרט תמיד מתקיים $X_i(U) \in \{X_{i-1}(U), X_{i-1}(U) + 1 - p, X_{i-1}(U) - p\}$, ולכן הם מקיימים את החסם הנדרש.

אחרי שהוכחנו את שתי התכונות של $X_0, \ldots, X_n$, כל שנותר הוא להפעיל את משפט אזומה לקבלת אי השוויון המבוקש $\Pr[|A \cap U| - p|A| > a] = \Pr[X_n > a] < e^{-a^2/2n}$.

צביעת גרף מושרה על קבוצה מקרית

לכל גרף $G = (V, E)$ ולכל קבוצת צמתים $V' \subset V$, נשים לב שמספרי הצביעה של הגרפים המושרים מקיימים $\chi(G) \le \chi(G[V']) + \chi(G[V - V'])$: מצביעה (חוקית) של $G[V']$ ב־k צבעים וצביעה של $G[V - V']$ ב־l צבעים קל לקבל צביעה של G ב־$k + l$ צבעים (צובעים את איברי V' בקבוצת צבעים זרה לזו שצובעים בה את צמתי $V - V'$). כאשר V' נבחר יוניפורמית מתקיים $\mathrm{E}[\chi(G[V'])] = \mathrm{E}[\chi(G[V - V'])]$, ולכן $2\mathrm{E}[\chi(G[V'])] \ge 1000$.

תהי עתה $c : V \to \{1, \ldots, 1000\}$ צביעה חוקית של G, ותהי $V_i = \{v \in V | c(v) = i\}$. נגדיר את מרטינגל החשיפה הבא: לכל $0 \le i \le 1000$, המ"מ X_i יציין את תוחלת מספר הצביעה של $G[V']$ כאשר כבר ידועים $V' \cap V_1, \ldots, V' \cap V_i$ (במילים אחרות, בשלב ה־i אנחנו חושפים את כל הבחירות מ־V_i). בפרט X_0 הוא המ"מ הקבוע $\mathrm{E}[\chi(G[V'])] \ge 500$, ו־$X_{1000}$ הוא המ"מ $\chi(G[V'])$ עצמו. נשים לב שהפונקציה $\chi(G[V'])$ מקיימת את תנאי ליפשיץ ביחס לחשיפות $V' \cap V_i$, כי כל V_i היא קבוצת צמתים ב"ת, ולכן שינוי ב־$V' \cap V_i$ לא משנה את מספר הצביעה ביותר מאחד. מכאן שאפשר להשתמש באי שוויון Azuma כדי לסיים את הטיעון:

$$\Pr[\chi(G[V']) < 400] \le \Pr[X_{1000} - X_0 < -\sqrt{10} \cdot \sqrt{1000}] < e^{-5} < \frac{1}{100}$$

(תרגיל זה נכתב במקור ע"י שריאל הר־פלד)

הפרדיגמה של פואסון

עוד על K_4 בגרף מקרי

נבחן את הגרף המקרי $G(n,\alpha n^{-2/3})$ עבור $\alpha>0$ קבוע נתון. הראו שקיים $\delta>0$ כך שלכל $n>3$ מתקיים שהסיכוי לכך שקיים K_4 בגרף הוא לפחות δ ולכל היותר $1-\delta$. מובן שמספיק להוכיח את זה עבור n גדול דיו.

פתרון לתרגיל על הפרדיגמה של פואסון

עוד על K_4 בגרף מקרי

נפתור את השאלה באמצעות אי־שוויון ינסון מתוך פרק הפרדיגמה של פואסון בתרגול. לכל רביעיית צמתים $1\leq i<j<k<l\leq n$ (כזכור קבוצת הצמתים של $G(n,p)$ היא $V=\{1,\ldots,n\}$), נסמן ב־$B_{i,j,k,l}$ את המאורע שאלו מהווים קליק בגרף. בסימונים של חוברת התרגול, הזוג $\{i,j,k,l\},\{i',j',k',l'\}$ מהווה קשת של גרף התלויות D אם ורק אם יש לשתי הקבוצות בין 2 ל־3 איברים משותפים. אצלנו $p_r=\alpha n^{-2/3}$ לכל $r\in V$. נוכל לחשב עתה $M=\prod_{1\leq i<j<k<l\leq n}\Pr[\neg B_{i,j,k,l}]=(1-\alpha^6n^{-4})^{\binom{n}{4}}$, וזה שואף ל־$e^{-\alpha^6/4!}$. בפרט קיים $\delta>0$ כך שעבור n גדול דיו מתקיים $\delta<M<1-2\delta$. כמו כן, בחישוב דומה לזה שעשינו עבור חסימת השונות בפרק על פונקציית סף לקיום K_4, מקבלים $\Delta=\binom{n}{2}\binom{n-2}{2}\binom{n-4}{2}\alpha^{11}n^{-22/3}+\binom{n}{3}(n-3)(n-4)\alpha^9n^{-6}=o(1)$. עבור n גדול דיו מתקיים גם $\Pr[B_{i,j,k,l}]<\frac{1}{2}$ (לכל i,j,k,l) וגם $e^{\Delta}<1+\delta$. לסיכום (עם הצבת $\epsilon=\frac{1}{2}$ באי־שוויון ינסון) מקבלים $\delta<M\leq\Pr[\bigwedge_{1\leq i<j<k<l\leq n}\neg B_{i,j,k,l}]\leq Me^{\Delta}<1-\delta$, שזה מה שרצינו להוכיח (זה אותו דבר אם מוכיחים את זוג החסמים לקיום קליק, או לאי־קיום קליק כפי שעשינו).

הלמה הלוקלית

צביעת קשתות בגרפים

הראו שלכל d קיים c כך שאם G הוא גרף מדרגה מקסימלית d (ומספר צמתים כל שהוא), אז ניתן לצבוע את הקשתות של G ב־c צבעים כך שבכל המעגלים הפשוטים בגרף יהיו קשתות משלושה צבעים לפחות ("מעגלים מאורך 2" אינם נחשבים).

קיום תת גרף ספציפי בגרף צפוף

הראו קיום קבוע גלובלי c עם המאפיין הבא: אם H הוא גרף בעל m צמתים ודרגה מקסימלית לכל היותר d, ו־G הוא גרף בעל $n>2^{cm}$ צמתים ולפחות $(\frac{1}{2}-\frac{1}{cd})n^2$ קשתות, אז G מכיל עותק של H כתת גרף (לא בהכרח מושרה).

הערה: יש משפט ידוע של Erdös ו־Stone על קיום תת גרף כזה עם תנאי אופטימלי על מספר הקשתות של G, אבל עם חסם רע בהרבה על n המינימלי, אשר משתמש בלמת הרגולריות על גרפים (שאותה לא נלמד בקורס זה).

שיפור קל של החסם על משפט רמזי

הראו שקיים גרף בעל $(\sqrt{2}/e-o(1))k2^{k/2}$ צמתים ושאין בו קליק או קבוצה בלתי תלויה בת k צמתים.

צביעות חסכוניות

הראו שגרף G עם דרגה מקסימלית Δ ניתן לצביעה (של הצמתים) ב־$\Theta(\Delta^{3/2})$ צבעים, כך שאין קשתות מופרות (מונוכרומטיות) ובנוסף לכך אין צומת עם שלושה שכנים מאותו צבע.

מילים לא חזרתיות

מילה $y \in \Sigma^{2m}$ נקראת חזרה אם קיימת מילה $x \in \Sigma^m$ כך ש־$y = xx$. מילה $w \in \Sigma^n$ נקראת חזרתית אם היא מכילה תת מילה רצופה שהיא חזרה, ואחרת היא נקראת לא חזרתית. הוכיחו כי מעל א"ב גדול דיו, קיימות מילים לא חזרתיות ארוכות כרצוננו. כלומר, הוכיחו כי קיים $k \in \mathbb{N}$, כך שלכל $n \in \mathbb{N}$ קיימת מילה לא חזרתית מעל א"ב בן k איברים שאורכה הוא לפחות n.

חלוקה בנטל

הראו שלכל k קיים קבוע C_k עם התכונה הבאה: נניח ש־$G = (V, E)$ הוא גרף מכוון, עם דרגת כניסה חסומה ע"י k (ז"א שאף צומת אינו צומת יציאה של יותר מ־k צמתים עם קשת אליו). ניתן אז לחלק את קבוצת הצמתים של G לשתי קבוצות V_1 ו־V_2, כך שגם בגרף המושרה על V_1 וגם בגרף המושרה על V_2, כל צומת v שדרגת היציאה שלו ב־G היתה $d(v) \geq C_k$, דרגת היציאה שלו בגרף המושרה המתאים תהיה בין $\frac{1}{3}d(v)$ לבין $\frac{2}{3}d(v)$.

פתרונות לתרגילים על הלמה הלוקלית

צביעת קשתות בגרפים

נגריל לכל קשת בגרף צבע מ־$\{1, \ldots, c\}$ באופן יוניפורמי וב"ת, ונוכיח שבהסתברות חיובית אין מעגל שמכיל פחות משלושה צבעים. לכל מעגל C בגרף נבחר באופן שרירותי שלוש קשתות עוקבות בו, ונסמן ב־A_C את המאורע ששלושת הקשתות לא צבועות בשלושה צבעים שונים. בנוסף לכך, אם לשני מעגלים C_1, C_2 בחרנו את אותן שלוש קשתות, אז נרשום את המאורע המתאים רק פעם אחת (נניח שרק ל־C_1, ואז נתעלם מ־C_2). מספיק להראות עתה שבהסתברות חיובית אף אחד מהמאורעות A_C לא יקרה.

לכל C מתקיים $\Pr[A_C] \leq 3c^{-1}$. בנוסף לכך, כל A_C הוא ב"ת בכל המאורעות אשר נרשמו עבור קשתות הזרות לקשתות שנבחרו מ־C. על כן A_C הוא ב"ת בכל המאורעות האחרים פרט ללא יותר מ־$9d^2$ מהם (יש 9 חפיפות־קשת אפשריות בין שני מסלולים מאורך 3, ומכיוון ש־d היא הדרגה המקסימלית יש לא יותר מעוד d^2 אפשרויות לבחור "קשתות המשך" למסלול החופף). בחירת של $c = 100d^2$ (למשל) תבטיח עתה שיתקיים $e \cdot 3c^{-1}(9d^2) < 1$, כך שנוכל להשתמש במקרה הסימטרי של הלמה הלוקלית ולסיים את ההוכחה.

קיום תת גרף ספציפי בגרף צפוף

ניתן להוכיח את המבוקש עם $c = 20$, ואת זאת נעשה עתה. אנו נגריל פונקציה f מ־$V(H)$, קבוצת הצמתים של H, לתוך $V(G)$, ע"י כך שלכל צומת u של H נבחר את $f(u)$ באופן מקרי וב"ת מ־$V(G)$ (בשביל ששאר הטיעון יעבוד, אי אפשר עדיין לבחור את f "בלי חזרות"). לכל קשת u, v של H נגדיר את המאורע E_{uv} בתור המאורע ש־$f(u), f(v)$ אינה קשת ב־G.

נשים לב עתה שמתקיים $\Pr[E_{uv}] < \frac{1}{10d}$, וכן שמאורע זה ב"ת בכל המאורעים האחרים (ז"א באלגברה הנוצרת על ידם) פרט לאלו הקשורים בקשתות של H המכילות את u או v. מהסוג האחרון יש פחות מ־$2d$ מאורעות, ולכן ניתן להפעיל את המקרה הסימטרי של הלמה הלוקלית ולקבל שבהסתברות חיובית אף מאורע לא קורה. אבל להמשך הטיעון צריך גם להשתמש בחסם (הקטן) שהלמה נותנת על ההסתברות, שהוא $(1 - \frac{1}{2d})^{md/2} > 2^{-m}$.

לסיכום, נחסום עתה את ההסתברות למאורע ש־f אינה חד־חד ערכית. לפי איחוד מאורעות (עם הנתון על n) זה חסום ע"י $\binom{m}{2}/n < 2^{-m}$. מכאן שבהסתברות חיובית גם f חד־חד ערכית וגם כל הקשתות של H עוברות לקשתות של G, ז"א שב־G חייב להיות עותק של H כתת־גרף.

שיפור קל של החסם על משפט רמזי

על מנת להראות את תוצאת השאלה, צריך להראות בעצם שלכל $C < \sqrt{2}/e$ ולכל k גדול דיו (כפונקציה של ההפרש בין C ו־$\sqrt{2}/e$), קיים גרף בעל $Ck2^{k/2}$ צמתים ושאין בו קליק או קבוצה ב"ת בת k צמתים.

נסתכל על הגרף $G(n, \frac{1}{2})$ עבור $n = Ck2^{k/2}$, ולכל קבוצת U בת k צמתים נגדיר את המאורע A_U שקבוצה זו מהווה קליק או קבוצה ב"ת. הסיכוי למאורע זה הוא $2^{1-\binom{k}{2}}$. כמו כן, כל מאורע A_U אינו תלוי בקבוצת כל המאורעות האחרים הנ"ל פרט ללא יותר מ־$\binom{k}{2}\binom{n}{k-2} - 1$ מתוכם (כי $\binom{k}{2}\binom{n}{k-2}$ הוא חסם על מספר הקבוצות מגודל k אשר חותכות את U ב־2 מקומות לפחות, ול־$k > 2$ המספר האמיתי קטן ממש מהחסם). אנו נרצה להשתמש בגרסה הסימטרית של הלמה הלוקלית על מנת להוכיח שבסיכוי חיובי אף אחד מהמאורעות הנ"ל אינו קורה, ולשם כך עלינו להוכיח שמתקיים $e2^{1-\binom{k}{2}}\binom{k}{2}\binom{n}{k-2} < 1$. שימוש בחסמים הידועים על בינומים ישלים את ההוכחה (שימו לב שהשוויון האחרון נכון רק תחת ההנחות על C):

$$\begin{aligned}
e2^{1-\binom{k}{2}}\binom{k}{2}\binom{n}{k-2} &< e2^{1+k/2-k^2/2} \cdot \frac{k^2-k}{2} \cdot (\frac{en}{k-2})^{k-2} \\
&= e2^{1+k/2-k^2/2} \cdot \frac{k^2-k}{2} \cdot (\frac{eCk}{k-2}2^{k/2})^{k-2} \\
&= (k^2-k)e^{k-1}2^{-k/2}C^{k-2}(1+\frac{2}{k-2})^{k-2} \\
&= O(k^2(\frac{eC}{\sqrt{2}})^k) = o(1)
\end{aligned}$$

ז"א שעבור k גדול דיו המכפלה הנ"ל אכן קטנה מ־1.

צביעות חסכוניות

נבצע צביעה אקראית של הגרף G ב־k צבעים (אחר כך נקבע את ערך k, אבל כבר נניח שהוא לפחות 3), ונגדיר את המאורעות ה"רעים" על מנת להשתמש בלמה הלוקלית. ישנם שני סוגים של מאורעות כאלו.

- לכל קשת $uv \in E$ נגדיר את המאורע A_{uv} שהיא מונוכרומטית. מתקיים $\Pr[A_{uv}] = \frac{1}{k}$.
- לכל שלושה צמתים u, v, w שיש עבורם שכן משותף, נגדיר את המאורע B_{uvw} שלשלושתם אותו צבע. מתקיים כי $\Pr[B_{uvw}] = \frac{1}{k^2}$.

אנו נראה את קיום תנאי הלמה הלוקלית בגרסא הלא־סימטרית ה"נוחה לשימוש" שבחוברת התרגול. ההסתברות של כל המאורעות קטנה מ־$\frac{1}{2}$ עבור $k \geq 3$, ועתה נבדוק לכל סוג מאורע את קיום התנאי השני.

מאורע A_{uv} מהסוג הראשון יהיה בפרט בלתי תלוי באלגברה הנוצרת ע"י כל המאורעות המתייחסים אך ורק לצבעים של הצמתים $V \setminus \{u\}$. יש לכל היותר $\Delta - 1$ מאורעות מהסוג הראשון שמתיחסים ל־u (לא כולל A_{uv} עצמו) ו־$\Delta\binom{\Delta-1}{2}$ מאורעות מהסוג השני שמתיחסים ל־u. לכן צריך להתקיים $\frac{1}{4} \leq \frac{\Delta(\Delta-1)(\Delta-2)}{2k^2} + \frac{\Delta-1}{k}$, וזה מתקיים בפרט לכל בחירה $k \geq \max\{2\Delta^{3/2}, 128\}$.

מאורע B_{uvw} מהסוג השני יהיה בלתי תלוי באלגברה הנוצרת ע"י כל המאורעות שאינם מתיחסים ל־u או v. יש לכל היותר 2Δ מאורעות מהסוג הראשון שמתיחסים ל־u או v, ולכל היותר $2\Delta\binom{\Delta-1}{2}$ מאורעות מהסוג השני (יש קצת ספירה כפולה). לכן צריך להתקיים $\frac{1}{4} \leq \frac{\Delta(\Delta-1)(\Delta-2)}{k^2} + \frac{2\Delta}{k}$, וזה יתקיים עבור $k \geq \max\{8\Delta^{3/2}, 64\}$.

סה"כ, על מנת שהלמה הלוקלית תתן לנו סיכוי חיובי שאף אחד מהמאורעות המוגדרים לא יתקיים (ואז הצביעה היא כנדרש) ניתן למשל להציב $k = \max\{8\Delta^{3/2}, 128\} = \Theta(\Delta^{3/2})$.

מילים לא חזרתיות

נקבע ערך $n \in \mathbb{N}$ כל שהוא, ונסמן ב־$A_{i,m}$ את המאורע שתת־המחרוזת $w_i, w_{i+1}, \ldots w_{i+2m-1}$ היא חזרה. בבירור מתקיים $\Pr[A_{i,m}] = k^{-m}$. כמו כן, המאורע $A_{i,m}$ בהכרח בלתי תלוי בכל המאורעות $A_{j,p}$ עבורם $\{i, i+1, \ldots, i+2m-1\} \cap \{j, j+1, \ldots, j+2p-1\} = \emptyset$. מספר הקטעים מאורך $2l$ שנחתכים עם הקטע שמתאים ל־$A_{i,m}$ הוא לכל היותר $2m+2l$. נסמן ב־$N_{i,m}$ את קבוצת (זוגות) האינדכסים שעבורם הקטעים המתאימים נחתכים עם $\{i, i+1, \ldots, i+2m-1\}$. אנו נשתמש בגרסה הכללית ביותר של הלמה הלוקלית. נקבע את $x_{i,m} = \frac{1}{6^m+1}$, ונראה שאכן תנאי הלמה מתקיימים עבור k גדול מספיק (בשורה השניה אנחנו משתמשים בזה שמתקיים $e^{-2x} < 1-x$ עבור $0 < x \le \frac{1}{2}$):

$$\begin{aligned}
\frac{1}{6^m+1} \prod_{(j,p)\in N_{i,m}} \left(1 - \frac{1}{6^p+1}\right) &\ge \frac{1}{6^m+1} \prod_{l=1}^{n/2} \left(1 - \frac{1}{6^l+1}\right)^{2m+2l} \\
&> \frac{1}{6^m+1} \prod_{l=1}^{n/2} \exp\left(-\frac{2}{6^l+1}\right)^{2m+2l} \\
&= \frac{1}{6^m+1} \exp\left(-\sum_{l=1}^{n/2} \frac{4m+4l}{6^l+1}\right) \\
&\ge \frac{1}{6^m+1} \exp\left(-4m \sum_{l=1}^{\infty} \frac{1}{6^l+1} - \sum_{l=1}^{\infty} \frac{4l}{6^l+1}\right) \\
&> \frac{1}{6^m+1} \exp\left(-m-4\right) > \exp\left(-2m-m-4\right) \ge \exp(-7m)
\end{aligned}$$

לכן אם $k \ge \lceil e^7 \rceil$ נקבל לכל i, m כי

$$x_{i,m} \prod_{(j,p)\in N_{i,m}} (1 - x_{j,p}) = \frac{1}{6^m+1} \prod_{A_{j,p}\in N_{i,m}} \left(1 - \frac{1}{6^p+1}\right) > k^{-m} = \Pr[A_{i,m}]$$

כנדרש (חישובים זהירים יותר יניבו חסם סביר יותר על k).

חלוקה בנטל

גם שאלה זו דורשת שימוש בניסוח הכללי ביותר של הלמה הלוקלית. לכל צומת $v \in V$ נגריל באופן מקרי, יוניפורמי וב"ת האם הוא ב־V_1 או ב־V_2. לכל v מדרגת יציאה $d(v) > C_k$ (אח"כ נקבע את C_k), נסמן ב־B_v את המאורע שהדרגה שלו בתת הגרף המתאים אינה בין $\frac{1}{3}d(v)$ לבין $\frac{2}{3}d(v)$. ל־v מדרגה שאינה עולה על C_k פשוט נגדיר את B_v כמאורע בסיכוי 0. לפי חסמי סטיות גדולות, קיים $\alpha > 0$ כך ש־$\Pr[B_v] < 2^{1-\alpha d(v)}$ לכל v.

עתה לכל v נגדיר את המספר $x_v = 1 - 2^{-\alpha/k}$, ואחרי זה נקבע את C_k להיות גדול דיו על מנת שיתקיים $2^{\alpha C_k/k} \ge 2/(1-2^{-\alpha/k})$. נסמן ב־$D_v$ את קבוצת המאורעות $\{B_w : \exists u(vu, wu \in E)\}$, ז"א את כל המאורעות הקשורים בצמתים שיש להם צומת יציאה משותף עם v. נשים לב ש־B_v הוא ב"ת לחלוטין במאורעות שאינם ברשימה D_v, וכן נשים לב שמתקיים $|D_v| \le (k-1) \cdot d(v)$ לפי הנתון על כך שכל דרגות הכניסה חסומות ע"י k.

על כן אם $d(v) \ge C_k$ אז $\prod_{w\in D_v}(1-x_w) \ge 2^{-\alpha(k-1)d(v)/k} \ge 2^{\alpha C_k/k} 2^{-\alpha d(v)} \ge 2^{1-\alpha d(v)}/(1-2^{-\alpha/k})$, ואז ניתן לוודא שמתקיים $\Pr[B_v] < 2^{1-\alpha d(v)} \le x_v \prod_{w\in D_v}(1-x_w)$, ז"א שאפשר להפעיל את הלמה הלוקלית ולראות שקיימת חלוקה עבורה אף מאורע B_v אינו מתקיים, כנדרש.

קורלציות

שוויון אצל קלייטמן

מצאו דוגמא (מעל S מתאים) שבה $\mathcal{A},\mathcal{B}\subset\mathcal{P}(S)$ הן משפחות מונוטוניות עולות, שתיהן אינן ריקות ואינן שוות ל־$\mathcal{P}(S)$, ומתקיים $|\mathcal{A}||\mathcal{B}|=2^{|S|}|\mathcal{A}\cap\mathcal{B}|$.

אי שוויון אצל קלייטמן

נתון ש־$\mathcal{A},\mathcal{B}\subset\mathcal{P}(S)$ הן משפחות מונוטוניות עולות לא ריקות, וכן ששתיהן כוללות אך ורק תתי קבוצות של S מגודל גדול מ־$\frac{1}{2}|S|$. הראו שבהכרח מתקיים $|\mathcal{A}||\mathcal{B}|<2^{|S|}|\mathcal{A}\cap\mathcal{B}|$.

נחתכים ולא מכסים

נתונה משפחה $\mathcal{F}$ של תתי־קבוצה של $S=\{1,\ldots,n\}$, כאשר $n\geq 2$. נתון גם שלכל $A,B\in\mathcal{F}$ מתקיים $A\cap B\neq\emptyset$ וכן $A\cup B\neq S$. הוכיחו ש־$\mathcal{F}$ מכילה לא יותר מ־2^{n-2} קבוצות. כמו כן תנו דוגמה ל־$\mathcal{F}$ אפשרית שזה הגודל שלה בדיוק.

חיתוך של מאורעות

א. הראו שאם $A_1,\ldots,A_k$ סדרה של תכונות מונוטוניות עולות של גרפים בעלי קבוצת הצמתים V (ז"א שאם $G(V,E)$ מקיים את A_i ו־$E\subset E'$ אז $G(V,E')$ גם מקיים את A_i), אז עבור גרף מקרי $G=G(n,p)$ מתקיים אי השוויון $\Pr[G\models A_1,\ldots,G\models A_n]\geq\prod_{i=1}^{k}\Pr[G\models A_i]$ (הסימון $G\models A$ פירושו לצורך העניין הוא ש־G מקיים את התכונה A).

ב. הוכיחו או הפריכו עבור $G(n,\frac{1}{2})$: (i) בהסתברות לפחות $1-2^{-\Omega(n^3)}$ הגרף G מכיל משולש. (ii) כאשר n הוא אי זוגי, בהסתברות לפחות 2^{-n} הדרגה המינימלית של הגרף היא לפחות $\frac{n-1}{2}$.

למתוח את הקשר

עבור הגרף המקרי $G(n,\ln(n)/n)$, הראו שבהסתברות $\Omega(1)$ (אבל שלא שואפת ל־1) הגרף יהיה קשיר.

לא מכילים

נניח ש־$\mathcal{F}\subset\mathcal{P}(\{1,\ldots,n\})\setminus\{\emptyset\}$ היא משפחה של תתי־קבוצה לא־ריקים של $\{1,\ldots,n\}$ שמתקיים עבורה $\sum_{F\in\mathcal{F}}2^{-|F|}\leq n/\log(n)$. הראו שקיימת קבוצה $B\subset\{1,\ldots,n\}$ שאינה מכילה אף איבר ב־$\mathcal{F}$ ומקיימת $|B|\geq(\frac{1}{2}-o(1))n$.

הבהרה: בכתיבה מפורשת (ללא שימוש ב־"o"), צריך להוכיח שלכל $\alpha<\frac{1}{2}$ קיים N_α, כך שאם $n>N_\alpha$ ו־$\mathcal{F}$ היא משפחה כמו בנוסח השאלה, אז קיימת B שאינה מכילה אף איבר ב־$\mathcal{F}$ ומקיימת $|B|\geq\alpha n$.

לרצות את כולם

נתונים מרחבי הסתברות $\mu_1,\ldots,\mu_k$, כולם מעל אותה קבוצת בסיס סופית S. נגריל מאורע E (כזכור מאורע הוא ת"ק של S) באופן יוניפורמי מבין כל תתי־הקבוצה של S. הראו שבהסתברות לפחות 2^{-k}, מתקיים $\mu_i(E)\geq\frac{1}{2}$ לכל $1\leq i\leq k$.

פונקציות מונוטוניות

תהי $f:\{0,1\}^n\to\{0,1\}$ פונקציה מונוטונית לא יורדת (אם $x,y\in\{0,1\}^n$ שני וקטורים המקיימים את אי השוויונות $x_1\geq y_1,\ldots,x_n\geq y_n$ אז מתקיים $f(x_1,\ldots,x_n)\geq f(y_1,\ldots,y_n)$). הראו שהפונקציה המונוטונית

לא עולה הקרובה ביותר ל־f במרחק האמינג (Hamming) היא פונקציה קבועה (ליתר דיוק שאחת מהפונקציות הנ"ל היא פונקציה קבועה, יש מקרים בהם זו אינה הפונקציה היחידה).

תזכורת: מרחק האמינג הלא־מנורמל הוא גודל הקבוצה $\{x \in \{0,1\}^n : f(x) \neq f(y)\}$. אם רוצים לחשב את המרחק המנורמל אז מחלקים ב־2^n, הגודל של תחום הפונקציות כאן.

משפט קלייטמן לרב־קבוצות

יהיו $\mathcal{A}, \mathcal{B}$ משפחות של רב־קבוצות מעל קבוצה S, כאשר כל איבר מ S יכול להופיע לכל היותר r פעמים באיברים של $\mathcal{A}$ או $\mathcal{B}$. נניח גם כי משפחות אלה מונוטוניות עולות, כאשר הכלה כוללת גם שמספר המופעים של איבר בקבוצה המכילה גדול או שווה לזה בקבוצה המוכלת. הראו שבמקרה זה $|\mathcal{A}|\,|\mathcal{B}| \leq (r+1)^{|S|}\,|\mathcal{A} \cap \mathcal{B}|$.

פתרונות לתרגילים על קורלציות

שוויון אצל קלייטמן

נבחר $S = \{1, \ldots, k\}$ עבור $k \geq 2$ כל שהוא. נבחר את $\mathcal{A}$ להיות המשפחה $\{A \subseteq S : 1 \in A\}$ ואת $\mathcal{B}$ להיות $\{A \subseteq S : 2 \in A\}$. חישוב ישיר מראה עתה שמתקיים $|\mathcal{A}||\mathcal{B}| = 2^{2k-2} = 2^{|S|}|\mathcal{A} \cap \mathcal{B}|$.

אי שוויון אצל קלייטמן

כזכור, בהוכחה של אי שוויון קלייטמן משתמשים במשפט ארבעת הפונקציות על מנת להראות שהמשפחות המונוטוניות מקיימות את אי השוויון $|\mathcal{A} \sqcup \mathcal{B}||\mathcal{A} \sqcap \mathcal{B}| \leq |\mathcal{A}||\mathcal{B}|$. כמו בהוכחה המקורית מתקיים עבור שתי המשפחות המונוטוניות העולות $\mathcal{A} \sqcup \mathcal{B} = \mathcal{A} \cap \mathcal{B}$, כאשר ידוע לנו גם שהחיתוך לא ריק (הוא כולל את S, כי המשפחות המקוריות לא היו ריקות ולכן כללו את S).

עבור $\mathcal{A} \sqcap \mathcal{B}$ ידוע לנו שהיא אינה מכילה את הקבוצה הריקה, מכיוון ש־$\mathcal{A}$ וגם $\mathcal{B}$ מכילות אך ורק קבוצות מגודל גדול מ־$\frac{1}{2}|S|$, וחיתוך של כל שתי קבוצות כאלו אינו ריק. על כן $|\mathcal{A} \sqcup \mathcal{B}||\mathcal{A} \sqcap \mathcal{B}| < 2^{|S|}|\mathcal{A} \cap \mathcal{B}|$, ובזאת סיימנו את ההוכחה.

נחתכים ולא מכסים

נגדיר שתי משפחות נוספות של תתי־קבוצה של $S = \{1, \ldots, n\}$. המשפחה $\mathcal{A}$ תוגדר כמשפחת כל הקבוצות המוכלות באיבר כל שהוא של $\mathcal{F}$, והמשפחה $\mathcal{B}$ תוגדר כמשפחת כל הקבוצות המכילות איבר כל שהוא של $\mathcal{F}$. בפרט חייב להתקיים $\mathcal{F} \subseteq \mathcal{A} \cap \mathcal{B}$ (אבל לא בהכרח מתקיים שוויון), כי כל איבר של $\mathcal{F}$ בפרט מוכל בעצמו, וכן מכיל את עצמו.

עתה נחסום את הגדלים של המשפחות החדשות. עבור תת־קבוצה A כל שהיא של S, לא ייתכן שגם $A \in \mathcal{A}$ וגם $S \setminus A \in \mathcal{A}$, מכיוון שאחרת היינו יכולים לקחת את האיבר של $\mathcal{F}$ שמכיל את A ואת האיבר של $\mathcal{F}$ שמכיל את $S \setminus A$, והאיחוד של שני אלו היה שווה ל־S כולה, בסתירה להנחות. על כן הגודל של $\mathcal{A}$ הוא לכל היותר 2^{n-1}, כי מכל זוג אפשרי של קבוצה והמשלימה שלה לכל היותר אחד מהם יהיה ב־$\mathcal{A}$.

באופן דומה, הגודל של $\mathcal{B}$ חסום ע"י 2^{n-1}, כי אם גם B וגם $S \setminus B$ יהיו ב־$\mathcal{B}$, אז ע"י לקיחת איברי $\mathcal{F}$ המוכלים באלו נקבל סתירה להנחה שאין ב־$\mathcal{F}$ זוג איברים עם חיתוך ריק.

לבסוף, נשים לב ש־$\mathcal{A}$ היא משפחה מונוטונית לא־עולה מעצם הגדרתה (היא מוגדרת כ"משפחת כל הקבוצות המוכלות במשהו"), בעוד ש־$\mathcal{B}$ היא מונוטונית לא־יורדת. מכאן שאפשר להשתמש במשפט קלייטמן ולקבל $2^n|\mathcal{A} \cap \mathcal{B}| \leq |\mathcal{A}||\mathcal{B}|$, ולכן $|\mathcal{F}| \leq |\mathcal{A} \cap \mathcal{B}| \leq 2^{n-2}$ כנדרש.

באשר לדוגמה ל־$\mathcal{F}$ מגודל 2^{n-2} בדיוק, אפשר לקחת את משפחת כל תתי־הקבוצה המכילות את האיבר "1" אולם אינן מכילות את האיבר "2".

חיתוך של מאורעות

א. מראים זאת באינדוקציה על k. עבור $k=1$ המשפט טריויאלי. כמו כן נשים לב שאם $A_1,\ldots,A_k$ הן תכונות מונוטוניות לא יורדות, אז גם $\bigwedge_{i=1}^{k} A_i$ היא תכונה כזו. לכן מתקיים עבור $k>1$ (ע"י שימוש כפי שנעשה בכיתה במשפט FKG ולאחריו שימוש בהנחת האינדוקציה)

$$\begin{aligned}
\Pr\left[G \vDash \bigwedge_{i=1}^{k} A_i\right] &= \Pr\left[G \vDash \left(\bigwedge_{i=1}^{k-1} A_i\right) \wedge A_k\right] \\
&\geq \Pr\left[G \vDash \left(\bigwedge_{i=1}^{k-1} A_i\right)\right] \cdot \Pr\left[G \vDash A_k\right] \\
&\geq \left(\prod_{i=1}^{k-1} \Pr\left[G \vDash A_i\right]\right) \cdot \Pr\left[G \vDash A_k\right] = \prod_{i=1}^{k} \Pr\left[G \vDash A_i\right]
\end{aligned}$$

כנדרש.

ב. (i) עבור $G\left(n,\frac{1}{2}\right)$ מתקבל בהסתברות $2^{-\binom{n}{2}}$ הגרף הריק, שבפרט אינו מכיל משולש, ולכן ההסתברות שהגרף לא מכיל משולש היא לפחות $2^{-\Theta\left(n^2\right)} > 2^{-\Omega\left(n^3\right)}$.

ב. (ii) נוכיח זאת על ידי שימוש בסעיף א. נגדיר עבור $i \in [n]$ את E_i להיות המאורע שדרגת הצומת i בגרף היא לפחות $\frac{n-1}{2}$. לכל i מתקיים $\Pr\left[E_i\right] \geq \frac{1}{2}$, וכל המאורעות הנ"ל מתייחסים לקיום תכונות מונוטוניות לא יורדות של הגרף, כך שמתקיימים התנאים הדרושים לשימוש בסעיף א.

למתוח את הקשר

גם כאן נחסום לכל קבוצה U את ההסתברות שלא יהיו קשתות בין U ל־V. עם זאת, על מנת לתרגם חסמים כאלו לחסם גלובלי על כל המאורעות (בעיקר עבור קבוצות בנות איבר בודד) לא מספיק יהיה להשתמש באיחוד מאורעות. תחת זאת נשתמש במשפט FKG כפי שמשתמשים בו בשאלה "חיתוך של מאורעות" מהפרק הזה. לכל $U \subset V$ שמקיים $1 \leq |U| \leq \frac{n}{2}$ נסמן ב־A_U את המאורע שיש לפחות קשת אחת בין U ל־$V \setminus U$, ואז ההסתברות שהגרף יהיה קשיר תהיה שווה להסתברות חיתוך כל המאורעות הנ"ל, שלפי משפט FKG חסומה מלמטה ע"י $\prod_{1\leq|U|\leq n/2}\Pr[A_U] = \prod_{k=1}^{\lfloor n/2\rfloor}\prod_{|U|=k}\Pr[A_U]$.

עבור U בודד מגודל k, מתקיים:

$$\begin{aligned}
\Pr[A_U] &= 1-(1-\frac{\ln(n)}{n})^{k(n-k)} \geq 1-e^{-k(n-k)\ln(n)/n} \\
&\geq \exp(-e^{-k(n-k)\ln(n)/n}-\frac{1}{2}e^{-2k(n-k)\ln(n)/n}) \geq \exp(-\frac{1}{2}e^{-k(n-k)\ln(n)/n})
\end{aligned}$$

על כן, ההסתברות של גרף מקרי המוגרל לפי $G(n,\frac{\ln(n)}{n})$ להיות קשיר היא לפחות

$$\prod_{k=1}^{\lfloor n/2\rfloor}\prod_{|U|=k}\Pr[A_U] \geq \exp(-\frac{1}{2}\sum_{k=1}^{\lfloor n/2\rfloor}\binom{n}{k}e^{-k(n-k)\ln(n)/n}) \geq \exp(-\frac{1}{2}\sum_{k=1}^{\lfloor n/2\rfloor}(\frac{en}{k})^k e^{-k(n-k)\ln(n)/n})$$

על מנת להשלים את השאלה, כל שנותר הוא להוכיח, עבור n גדול מספיק, חסם עליון גלובלי קבוע על הסכום $\sum_{k=1}^{\lfloor n/2\rfloor}(\frac{en}{k})^k e^{-k(n-k)\ln(n)/n} = \sum_{k=1}^{\lfloor n/2\rfloor}\exp(k+\frac{k^2\ln(n)}{n}-k\ln k)$. נפצל את הטיפול לכל k קבוע שמקיים $1 \leq k < K$ עבור K מתאים (למשל $K = \lceil e^6\rceil$ יספיק לנו כאן), ולשאר הסכום.

עבור כל $1 \leq k < K$ קבוע, פשוט נשים לב שהביטוי $\exp(k+\frac{k^2\ln(n)}{n}-k\ln k)$ שואף לקבוע, ולכן בפרט יהיה לנו חסם עליון קבוע על $\sum_{k=1}^{K-1}\exp(k+\frac{k^2\ln(n)}{n}-k\ln k)$.

עבור $K \leq k \leq \lfloor \frac{n}{2}\rfloor$, ראשית נטען שעבור n גדול מקבוע מתאים (לא תלוי k) מתקיים $\frac{k\ln(n)}{n} \leq \frac{2\ln(k)}{3}$. לשם כך נחפש מקסימום של הביטוי $\frac{k\ln(n)}{n}-\frac{2\ln(k)}{3}$: נגזור אותו לפי k פעמיים ונקבל את הביטוי $2/3k^2$ שהוא חיובי בתחום הרלוונטי, ולכן המקסימום יכול להתקבל רק בקצוות, $k=K$ ו־$k=\lfloor\frac{n}{2}\rfloor$. כאשר n גדול דיו הביטוי $\frac{K\ln(n)}{n}-\frac{2\ln(K)}{3}$ הוא שלילי, בעוד שמתקיים $\frac{\lfloor n/2\rfloor\ln(n)}{n}-\frac{2\ln(\lfloor n/2\rfloor)}{3} \leq \frac{\ln(n)}{2}-\frac{2\ln(n)}{3}+1 \leq 0$, ובזאת הוכחנו את החסם.

מחסם זה נובע $\sum_{k=K}^{\lfloor n/2\rfloor}\exp(k+\frac{k^2\ln(n)}{n}-k\ln(k)) \leq \sum_{k=K}^{\lfloor n/2\rfloor}\exp(k-\frac{1}{3}k\ln(k)) \leq \sum_{k=K}^{\infty}e^{-k\ln(k)/6}$, והסכום הימני שואף לקבוע. בזאת הושלם החלק האחרון בהוכחה.

הערות: גם כאן יש שיטות אלטרנטיביות לעשות את החסמים, כולל שיטות "היברידיות" שמשתמשות בחסם יותר פשוט של איחוד מאורעות עבור k יותר גדולים ובקורלציות רק עבור ה־k הקטנים ביותר (או אפילו רק $k=1$, ז"א בנוגע לסיכוי שקיימים צמתים מבודדים). ידוע (עם הוכחה יותר מדוקדקת) שההסתברות לקבל גרף קשיר בהגרלה לפי $G(n,\frac{\ln(n)}{n})$ שואפת ל־e^{-1}.

לא מכילים

נקבע מספר ממשי $0<\alpha<\frac{1}{2}$ כל שהוא. נגריל את הקבוצה B באופן יוניפורמי מכל תתי־הקבוצה של $\{1,\ldots,n\}$ (ז"א שכל $i\in\{1,\ldots,n\}$ יבחר להיות ב־B בהסתברות $\frac{1}{2}$ באופן ב"ת בבחירות האחרות), ונראה, עבור n גדול דיו, ש־B לא תכיל אף אחד מהאיברים של $\mathcal{F}$ וגם תקיים $|B|\geq\alpha n$.

לכל $F\in\mathcal{F}$ נסמן ב־A_F את המאורע ש־B אינה מכילה את F. נשים לב שהמאורע הנ"ל (אם מסתכלים עליו כעל משפחת כל תתי־הקבוצה של $\{1,\ldots,n\}$ שאינם מכילים את F) הוא מונוטוני־לא עולה (כמשפחה של קבוצות). כמו כן מתקיים $\Pr[A_F]=1-2^{-|F|}$ ישירות מההגדרות. על כן ניתן להשתמש במשפט FKG (או לחילופין במשפט קלייטמן), ולקבל $\Pr[\bigwedge_{F\in\mathcal{F}}A_F]\geq\prod_{F\in\mathcal{F}}\Pr[A_F]=\prod_{F\in\mathcal{F}}(1-2^{-|F|})$ (כרגיל השתמשנו באינדוקציה עבור אי השוויון לסדרות של יותר מ־2 מאורעות).

עבור המשך הפיתוח, נשתמש בכך שמתקיים $(1-x)>e^{-2x}$ עבור $0<x<\frac{1}{2}$ (אפשרות הוכחה – הפונקציה e^{-2x} קמורה ממש ולכן חותכת את הישר $y=1-x$ בלא יותר משתי נקודות; אחת הנקודות הנ"ל היא $x=0$, והשניה נמצאת מימין ל־$x=\frac{1}{2}$). על כן מתקיים $\Pr[\bigwedge_{F\in\mathcal{F}}A_F]>e^{-2\sum_{F\in\mathcal{F}}2^{-|F|}}\geq e^{-2n/\log(n)}$.

מחסימת סטיות גדולות $\Pr[|B|<\alpha n]<e^{-2(1/2-\alpha)^2 n}$. מכיוון שעבור α קבוע $e^{-2(1/2-\alpha)^2 n}=o\big(e^{-2n/\log(n)}\big)$, עבור n גדול דיו הסתברות איחוד המאורעות ה"רעים" $|B|<\alpha n$ ו־$\neg\bigwedge_{F\in\mathcal{F}}A_F$ תהיה קטנה מ־1, ולכן בהסתברות חיובית יתקיים גם $|B|\geq\alpha n$ וגם $\bigwedge_{F\in\mathcal{F}}A_F$.

לרצות את כולם

השאלה מדברת על מרחבי הסתברות מרובים, אולם עיקר הפתרון הוא הזיהוי של מרחב ההסתברות שאנחנו מנתחים. קבוצת הבסיס שלנו תהיה קבוצת המאורעות האפשריים מעל S, ז"א $\mathcal{P}(S)$, והבחירה תהיה יוניפורמית מתוכה. כל משפחה $\mathcal{A}_i\subseteq\mathcal{P}(S)$ שמתאימה למאורע "$\mu_i(E)\geq\frac{1}{2}$" היא מונוטונית לא־יורדת.

משפט קלייטמן, במושגים הסתברותיים, אומר לנו שלכל שתי משפחות מונוטוניות לא־יורדות $\mathcal{A}$ ו־$\mathcal{B}$ מתקיים $\Pr[E \in \mathcal{A} \wedge E \in \mathcal{B}] \geq \Pr[E \in \mathcal{A}]\Pr[E \in \mathcal{B}]$. מכאן אפשר להוכיח באינדוקציה שעבור k המשפחות מתקיים $\Pr[\bigwedge_{i=1}^{k} E \in \mathcal{A}_i] \geq \Pr[E \in \mathcal{A}_1]\Pr[\bigwedge_{i=2}^{k} E \in \mathcal{A}_i] \geq \cdots \geq \prod_{i=1}^{k} \Pr[E \in \mathcal{A}_i]$.

לבסוף, מכיוון שלכל קבוצה E תמיד מתקיים $\mu_i(E) \geq \frac{1}{2}$ או $\mu_i(S \backslash E) \geq \frac{1}{2}$ (או שניהם), מתקיים $\Pr[E \in \mathcal{A}_i] \geq \frac{1}{2}$ לכל $1 \leq i \leq k$. לכן מאי־השוויון למעלה מתקיים $\Pr[\bigwedge_{i=1}^{k} E \in \mathcal{A}_i] \geq 2^{-k}$ כנדרש.

פונקציות מונוטוניות

נסמן ב־$\mathcal{F}_0 = \{x \in \{0,1\}^n : f(x) = 0\}$ את קבוצת האיברים שעליהם f מקבלת את הערך 0, ונסמן ב־$\mathcal{F}_1 = \{0,1\}^n \backslash \mathcal{F}_0$ את האיברים עליהם f היא 1. אם מזהים כל איבר $x \in \{0,1\}^n$ עם קבוצת האחדות שלו $S_x = \{i : x_i = 1\} \subseteq [n]$, אז קל לוודא ש־$\mathcal{F}_0$ היא משפחה מונוטונית לא עולה של תת קבוצות של $[n]$, וש־$\mathcal{F}_1$ היא משפחה מונוטונית לא יורדת. נניח עתה ש־$g : \{0,1\}^n \to \{0,1\}$ היא פונקציה מונוטונית לא עולה, ונסמן בדומה את המשפחות $\mathcal{G}_0$ (שהיא מונוטונית לא יורדת) ו־$\mathcal{G}_1$ (שהיא מונוטונית לא עולה).

המרחק של f מהפונקציה הקבועה הקרובה ביותר הוא $\min\{|\mathcal{F}_0|, |\mathcal{F}_1|\}$ שכן $|\mathcal{F}_0|$ הוא המרחק של f מהפונקציה הקבועה 1 ו־$|\mathcal{F}_1|$ הוא המרחק מהפונקציה הקבועה 0. כמו כן, המרחק מ־f ל־g הוא $|\mathcal{F}_0 \cap \mathcal{G}_1| + |\mathcal{F}_1 \cap \mathcal{G}_0|$. לכן על מנת להוכיח את טענת השאלה עלינו להוכיח שלכל g מונוטונית לא עולה מתקיים האי שוויון הבא: $|\mathcal{F}_0 \cap \mathcal{G}_1| + |\mathcal{F}_1 \cap \mathcal{G}_0| \geq \min\{|\mathcal{F}_0|, |\mathcal{F}_1|\}$.

לפי המשפט של קלייטמן מתקיים $2^n |\mathcal{F}_1 \cap \mathcal{G}_0| \geq |\mathcal{F}_1| \cdot |\mathcal{G}_0|$ ו־$2^n |\mathcal{F}_0 \cap \mathcal{G}_1| \geq |\mathcal{F}_0| \cdot |\mathcal{G}_1|$, ולכן:

$$
\begin{aligned}
|\mathcal{F}_0 \cap \mathcal{G}_1| + |\mathcal{F}_1 \cap \mathcal{G}_0| &\geq 2^{-n}\left(|\mathcal{F}_1| \cdot |\mathcal{G}_0| + |\mathcal{F}_0| \cdot |\mathcal{G}_1|\right) \\
&= 2^{-n}\left(|\mathcal{F}_1| \cdot (2^n - |\mathcal{G}_1|) + |\mathcal{F}_0| \cdot |\mathcal{G}_1|\right) \\
&= \left(2^{-n}|\mathcal{G}_1|\right) \cdot |\mathcal{F}_0| + \left(1 - 2^{-n}|\mathcal{G}_1|\right) \cdot |\mathcal{F}_1| \\
&\geq \left(2^{-n}|\mathcal{G}_1|\right) \cdot \min\{|\mathcal{F}_0|, |\mathcal{F}_1|\} + \left(1 - 2^{-n}|\mathcal{G}_1|\right) \cdot \min\{|\mathcal{F}_0|, |\mathcal{F}_1|\} \\
&= \min\{|\mathcal{F}_0|, |\mathcal{F}_1|\}
\end{aligned}
$$

נשים לב שכאמור אכן יש מקרים שקיימת פונקציה מונוטונית לא עולה קרובה ביותר שאינה קבועה. נביט לדוגמא על המקרה $n = 2$ ועל הפונקציה המונוטונית לא יורדת $f(0,0) = f(1,0) = 0, f(0,1) = f(1,1) = 1$. מכוון שהפונקציה הקבועה הקרובה ביותר היא הפונקציה הזהותית 1 או הזהותית 0, המרחק מהפונקציה המונוטונית לא עולה הקרובה ביותר הוא 2. עם זאת, ניתן להשיג אותו גם עם הפונקציה הלא־קבועה המוגדרת באופן הבא: $g(0,0) = g(0,1) = 1, g(1,0) = g(1,1) = 0$.

משפט קלייטמן לרב־קבוצות

נוכיח זאת בעזרת משפט FKG עבור הקבוצה $S' = S \times \{1, \ldots, r\}$. בהנתן רב־קבוצה C מעל S עם r עותקים לכל היותר מכל איבר, נאמר ש־$C' \subseteq S'$ מייצגת את C אם היא מורכבת בדיוק מכל האיברים מהצורה (a, i) כאשר a מופיע ב־C i פעמים או יותר. מיפוי זה הוא חד חד ערכי ומשמר הכלה. נגדיר פונקציה $\delta : \mathcal{P}(S') \to \mathbb{R}^+$ שמציינת את תתי הקבוצות ב־S' שמייצגות רב־קבוצות מעל S (כלומר $\delta(C') = 1$ אם C' מייצגת רב־קבוצה C כלשהי, ואחרת $\delta(C') = 0$). אם שתי קבוצות A, B מייצגות רב־קבוצות, אז כך גם $A \cap B$ ו־$A \cup B$, ולכן δ היא לוג־סופר־מודולרית.

נגדיר את f המציינת קבוצות C המכילות קבוצה D המייצגת רב־קבוצה מ־$\mathcal{A}$ ובאותו אופן נגדיר את g עבור $\mathcal{B}$. כיוון ששתי הקבוצות מונוטוניות עולות והמעבר לקבוצה מייצגת משמר הכלה, אז אם $f(C) = 1$ ו־C מייצגת איזה רב־קבוצה, אז אכן רב־קבוצה זו חברה ב־$\mathcal{A}$, ובאותו אופן עבור g ו־$\mathcal{B}$.

כעת נציב באי שוויון FKG:

$$\left(\sum_{C\subseteq S'} f\left(C\right)\delta\left(C\right)\right)\left(\sum_{C\subseteq S'} g\left(C\right)\delta\left(C\right)\right) \quad\leq\quad \left(\sum_{C\subseteq S'} f\left(C\right)g\left(C\right)\delta\left(C\right)\right)\left(\sum_{C\subseteq S'}\delta\left(C\right)\right)$$

כעת, נשים לב ש $\delta\left(C\right)\cdot f\left(C\right)=1$ אם ורק אם C מייצגת רב־קבוצה מ־$\mathcal{A}$, ובאופן דומה עבור g ו־$\mathcal{B}$, ועבור $f\cdot g$ ו־$\mathcal{A}\cap\mathcal{B}$. לסיכום, לפי מספרן האפשרי של רב־קבוצות מעל S ידוע לנו כי $\sum_{C\subseteq S'}\delta\left(C\right)=\left(r+1\right)^{|S|}$, ובכך מסתיימת ההוכחה.

אנטרופיה

לא ממריא

נתון ש־X הוא משתנה מקרי מעל קבוצת כל המספרים הטבעיים (לא סופית, אבל עדיין בדידה), ונתון שהתוחלת של X היא סופית. הראו שגם האנטרופיה של X היא בהכרח סופית.

הטלות נחתכות

תהא $\mathcal{F}$ משפחה של וקטורים ב־$S_1\times S_2\times\cdots\times S_n$ ויהי $\mathcal{G}=\{G_1,G_2,\ldots,G_m\}$ אוסף של תת קבוצות של $[n]$ כך שכל איבר $i\in[n]$ מופיע בלפחות k קבוצות ב־$\mathcal{G}$. עבור $1\leq i\leq m$ נסמן ב־$\mathcal{F}_i$ את הקבוצה הנוצרת על ידי הטלת כל איברי $\mathcal{F}$ על הקואורדינטות ב־G_i. הוכיחו שאז מתקיים $|\mathcal{F}|^k\leq\prod_{i=1}^m|\mathcal{F}_i|$.

החשובים הנחשבים

עבור מרחב הסתברות μ מעל קבוצה סופית S ועבור $\epsilon>0$ כל שהוא, הראו שקיימת תת־קבוצה $S'\subseteq S$ שמקיימת $\Pr_\mu[S']\geq 1-\epsilon$ וכן $|S'|\leq 2^{\mathrm{H}[\mu]/\epsilon}$.

שני שימושים באי־שוויון פינסקר

אי־שוויון פינסקר (Pinsker) קובע את הדבר הבא: עבור שני מרחבי הסתברות μ ו־ν מעל אותה קבוצת בסיס בדידה S, מתקיים $d(\mu,\nu)\leq\sqrt{\frac{1}{2}D(\mu\|\nu)}$, כאשר d מסמן את המרחק בין ההתפלגויות (ראו את הפרק הראשון בחוברת זו) ו־D מסמן את הפיצוליות (אנטרופיה יחסית – Kullback-Leibler divergence).

א. נתון ש־X ו־Y הם שני משתנים מקריים מעל מרחב הסתברות בדיד, שמקבלים את כל ערכיהם בטווח הממשיים $[0,1]$. השתמשו באי השוויון למעלה על מנת להראות שמתקיים $|\mathrm{Cov}[X,Y]|\leq\sqrt{\frac{1}{2}\mathrm{I}[X,Y]}$.

ב. נתון מרחב הסתברות μ מעל קבוצה סופית S. בנוסף, נסמן ב־π את מרחב ההסתברות היוניפורמי מעל S, וב־$d(\mu,\pi)$ את המרחק בין שני מרחבי ההסתברות. הראו שמתקיים $\mathrm{H}[\mu]\leq\log(|S|)-2(d(\mu,\pi))^2$.

סאב־מודולריות אנטרופית

נניח ש־$X_1,\ldots,X_n$ הם משתנים מקריים, כאשר X_i מקבל ערכים מהקבוצה S_i, ולכל $A\subseteq\{1,\ldots,n\}$ נגדיר מ"מ X_A שמקבל ערכים מהקבוצה $\prod_{i\in A}S_i$ לפי $X_A=\langle X_i:i\in A\rangle$ (ז"א ש־X_A הוא סדרת הערכים המתאימים ל־X_i עבור $i\in A$). נגדיר פונקציה $\eta:\mathcal{P}(\{1,\ldots,n\})\to\mathbb{R}$ לפי $\eta(A)=\mathrm{H}[X_A]$.

הראו שפונקציה זו היא סאב־מודולרית: לכל A,B מתקיים $\eta(A)+\eta(B)\geq\eta(A\cup B)+\eta(A\cap B)$.

הערה: אפשר ככה לתת הוכחה אלטרנטיבית לאי־שוויון שירר מהתרגול. נסו להראות שלכל פונקציה סאב־מודולרית η ולכל משפחה $\mathcal{A}$ של קבוצות כך שכל i נמצא בלפחות k מהן, מתקיים $k\eta(\{1,\ldots,n\}) \leq \sum_{A\in\mathcal{A}} \eta(A)$.

רעש רקע

נתונים שלושה מ"מ בדידים X,Y,Z מעל מרחב הסתברות כל שהוא, שמקבלים ערכים ב־$\mathbb{R}$. נתון ש־X הוא ב"ת בצירוף של Y ו־Z (לא רק בכל אחד מהם לחוד). לא נתון האם Y ו־Z הם ב"ת זה בזה או לא. הראו שמתקיים $\mathrm{H}[X+Y|Z] \geq \mathrm{H}[Y|Z]$ וכן $\mathrm{H}[Y|X+Z] \geq \mathrm{H}[Y|Z]$. שימו לב שאלו פעולות חיבור של מספרים ממשיים ולא פעולות שרשור משתנים.

הערה: לא קשה למצוא דוגמאות שבהן אי־השוויונים האלו אינם מתקיימים אם מרשים ל־X להיות תלוי במשתנים האחרים. נסו לחשוב על כאלה.

לא תרבו

נתונה קבוצת מילים $C \subseteq \{0,1\}^n$. עבור $\alpha > 0$ קבוע, נתון שלכל קבוצת אינדכסים $I \subseteq \{1,\ldots,n\}$ המקיימת $|I| \leq \alpha n$ קיימת מילה $w = w_1,\ldots,w_n \in C$ המתאפסת על כל האינדכסים ב־I, ז"א $w_i = 0$ לכל $i \in I$. כמו כן, נתון שקיים מרחב הסתברות μ מעל $\{1,\ldots,n\}$, כך שאם i הוא אינדכס הנבחר לפי מרחב הסתברות זה, אז לכל $w = w_1,\ldots,w_n \in C$ מתקיים $\Pr_{i\sim\mu}[w_i = 1] \geq \frac{9}{10}$. הראו שקיים $\beta > 0$ התלוי ב־α בלבד, שעבורו מתקיים $|C| \leq 2^{(1-\beta)n}$.

דומה בדומה

נתון ש־μ ו־ν הם שני מרחבי הסתברות מעל $S = \{1,\ldots,n\}$. הראו שההפרש בין האנטרופיות שלהם מקיים $|\mathrm{H}[\mu] - \mathrm{H}[\nu]| \leq d(\mu,\nu)\log(n-1) + H(d(\mu,\nu))$.

תזכורות: הסימון $d(\mu,\nu)$ הוא המרחק בין ההתפלגויות המופיע בתחילת חוברת זו. הפונקציה H בצד ימין היא פונקצית האנטרופיה $H(p) = p\log(\frac{1}{p}) + (1-p)\log(\frac{1}{1-p})$.

סודות ושקרים

חישבו על האפשרות לכתוב אלגוריתם דטרמיניסטי שנדרש למצוא ערך לא ידוע $k \in \{1,\ldots,n\}$ באמצעות q שלבים. בשלב ה־i של האלגוריתם, האלגוריתם בונה קבוצה A_i ומקבל תשובה לשאלה "האם $k \in A_i$". מותר לאלגוריתם לבנות את A_i בהסתמך על התשובות הקודמות עבור $A_1,\ldots,A_{i-1}$. לאחר q שלבים האלגוריתם פולט מספר $k' \in \{1,\ldots,n\}$.

הבעיה כאן היא שבכל שלב האלגוריתם מקבל את התשובה הנכונה לשאלה "האם $k \in A_i$" בהסתברות $\frac{9}{10}$, ומקבל תשובה שקרית בהסתברות $\frac{1}{10}$, כאשר ההסתברות לשקר היא ב"ת ביחס לכל מה שקרה בשלבים הקודמים של האלגוריתם (וגם ב"ת בזהות של A_i שהאלגוריתם בנה).

הראו שקיימים קבועים $\alpha,\beta > 0$ (לא תלויים ב־n), כך שעבור כל n גדול מספיק, אם האלגוריתם בסוף נותן את התשובה הנכונה (ז"א מתקיים $k' = k$) בהסתברות לפחות $1-\alpha$, אז בהכרח מתקיים $q \geq (1+\beta)\log(n)$.

הבהרות: לא נכתבו הגבלות על ה"אלגוריתם", ובאמת לא נתון עבורו זמן חישוב מסויים, או אפילו שהוא ניתן בכלל לחישוב. נתון רק שבניית A_i תלויה באופן דטרמיניסטי בתשובות שניתנו בשלבים הקודמים. אתם אמורים לתת חסם תחתון על מספר השלבים האפשרי של אלגוריתם "מוצלח" כזה: אם לכל $k \in \{1,\ldots,n\}$ מתקיים $k' = k$ בהסתברות מספיק גבוהה, אז מספר השלבים לא יכול להיות קטן.

הדרכה: מכיוון שהאלגוריתם הוא דטרמיניסטי, התשובה שלו תלויה אך ורק בסדרת התשובות ("כן" או "לא") שניתנו במהלך q השלבים (נהוג לכתוב אלגוריתם כזה בצורה של עץ־החלטות). כתבו משתנה מקרי עבור סדרה

זו, ונסו לחסום את האנטרופיה שלו ביחס למ"מ אחרים כאשר הערך k נבחר באופן יוניפורמי מתוך $\{1,\dots,n\}$.

גן השבילים המתפצלים

נתון גרף G עם n צמתים, שכל הצמתים שלו מדרגה 2 לפחות, ועם דרגה ממוצעת $d=\frac{1}{n}\sum_{v\in V}d(v)$. כמו כן נתון שגודל המעגל הפשוט המינימלי בו הוא לפחות $2k+1$. הראו שמתקיים $d(d-1)^{k-1}\le n$.

שימו לב: פתרון השאלה משתמש בתוצאת השאלה "לא חוזרים לאחור" מפרק השאלות על הילוכים מקריים.

משפחות נחתכות במשולשים

תהא $\mathcal{F}$ משפחת גרפים על קבוצת הצמתים $\{1,2,\dots,t\}$ כך שלכל שני גרפים ב־$\mathcal{F}$ ישנו משולש המופיע בשניהם. הוכיחו שאז מתקיים $|\mathcal{F}|<\frac{1}{4}2^{\binom{t}{2}}$.

משפחות נחתכות בשידוכים

תהא $\mathcal{F}$ משפחת גרפים על הצמתים $\{1,2,\dots,2n\}$ כך שלכל שני גרפים ב־$\mathcal{F}$ ישנו שידוך מושלם המופיע בשניהם. הראו כי $|\mathcal{F}|\le 2^{\binom{2n}{2}-n}$.

פתרונות לתרגילים על אנטרופיה

לא ממריא

נראה כאן שתי אפשרויות לפתרון השאלה.

פתרון ראשון: נסמן ב־a את התוחלת של X, ונניח $a>0$ (אחרת $\mathrm{H}[X]=0$ וסיימנו). זה נוח (אם כי אפשר גם בלי) להגדיר את המשתנה המקרי הנוסף הבא: $Y=\max\{0,\lfloor\log(X/a)\rfloor\}$. במילים אחרות, Y הוא המספר הטבעי המינימלי k שעבורו מתקיים $X\le a2^Y$. מכיוון ש־Y הוא פונקציה של X, מתקיים עבורם השוויון $\mathrm{H}[X]=\mathrm{H}[X,Y]=\mathrm{H}[Y]+\mathrm{H}[X|Y]$. עתה נחסום את שני המחוברים הימניים.

העיקר הוא לשים לב שלפי אי־שוויון מרקוב מתקיים $\Pr[Y=k]\le\Pr[X>a2^{k-1}]\le 2^{1-k}$. מכאן חוסמים את המחובר הראשון לפי $\mathrm{H}[Y]=\sum_{k=0}^{\infty}\Pr[Y=k]\log\frac{1}{\Pr[Y=k]}\le 3+\sum_{k=3}^{\infty}k2^{1-k}\le 6$ (עבור $0<x<\frac{1}{e}$ הפונקציה $x\log\frac{1}{x}=(x\ln\frac{1}{x})/\ln 2$ עולה לפי גזירה, ומפתחים $\sum_{k=1}^{\infty}k2^{-k}=\sum_{i=1}^{\infty}(\sum_{j=i}^{\infty}2^{-j})=2$).

עבור חסימת המחובר השני, ניזכר בהגדרה $\mathrm{H}[X|Y]=\sum_{k=0}^{\infty}\Pr[Y=k]\mathrm{H}[X|Y=k]$. ניתן לחסום כל מחובר לחוד ע"י $\Pr[Y=k]\mathrm{H}[X|Y=k]\le 2^{1-k}\log(a2^k+1)\le 2^{1-k}(\log(a)+1+k)$ (השתמשנו בכך ש־X יכול לקבל לא יותר מ־$a2^k+1$ ערכים שונים כאשר מתנים על $Y=k$). לכן יש לנו את החסם $\mathrm{H}[X|Y]\le\sum_{k=0}^{\infty}2^{1-k}(\log(a)+1+k)\le 4\log(a)+8$, וסה"כ $\mathrm{H}[Y]\le 4\log(a)+14$ (כמובן שממש לא ניסינו לתת כאן את הביטוי הכי טוב האפשרי).

פתרון שני: לשם הנוחות לכל $i\in\mathbb{N}$ נסמן $p_i=\Pr[X=i]$. נחלק את קבוצת המספרים הטבעיים לפי ההסתברויות שלהם לשתי הקבוצות $A=\{i\in\mathbb{N}:p_i\le 2^{-i}\}$ ו־$B=\{i\in\mathbb{N}:p_i>2^{-i}\}$. מתקיים השוויון $\mathrm{H}[X]=\sum_{i\in A}p_i\log\frac{1}{p_i}+\sum_{i\in B}p_i\log\frac{1}{p_i}$, ונחסום כל מחובר בנפרד.

עבור המחובר הראשון, מתקיים $\sum_{i\in A}p_i\log\frac{1}{p_i}\le 3+\sum_{i=2}^{\infty}i\cdot 2^{-i}\le 5$ (גם כאן השתמשנו במונוטוניות של $x\log\frac{1}{x}$ עבור $0<x<\frac{1}{e}$). עבור המחובר השני, נשתמש ב־$\sum_{i\in B}p_i\log\frac{1}{p_i}\le\sum_{i\in B}i\cdot p_i\le\mathrm{E}[X]$.

הטלות נחתכות

נסמן ב־$X = (X_1, \ldots, X_n)$ משתנה מקרי המקבל ערך מ־$\mathcal{F}$ בהתפלגות אחידה. נסמן ב־$X\left(G_i\right)$ את הטלת X על הקואורדינטות ב־G_i. לפי אי שוויון Shearer שנלמד בתרגול ידוע כי $k\mathrm{H}\left[X\right] \leq \sum_{i=1}^{m} \mathrm{H}\left[X\left(G_i\right)\right]$. מכוון ש־$X$ נבחר בהתפלגות אחידה אז $\mathrm{H}\left[X\right] = \log|\mathcal{F}|$, ומכוון ש־$X\left[G_i\right]$ הוא בחירה בתוך $\mathcal{F}_i$ אז $\mathrm{H}\left[X\left(G_i\right)\right] \leq \log|\mathcal{F}_i|$. כך מקבלים $k\log|\mathcal{F}| \leq \sum_{i=1}^{m} \log|\mathcal{F}_i|$, ובלקיחת 2 בחזקת שני צידי אי השוויון מקבלים את הטענה המבוקשת.

החשובים הנחשבים

נגדיר מ"מ $X : S \to \mathbb{R}$ לפי $X(s) = \log(1/\mu(s))$ לכל $s \in S$ (כאשר במקרה של $\mu(s) = 0$ נגדיר את $X(s)$ להיות ערך שרירותי גבוה מספיק, במקרה שלנו גבוה מ־$\mathrm{H}[\mu]/\epsilon$). כזכור מהגדרת האנטרופיה בשיעור מתקיים $\mathrm{H}[\mu] = \mathrm{E}_{s\sim\mu}[X(s)]$, ולכן מאי־שוויון מרקוב מתקיים $\Pr_{s\sim\mu}[X(s) \geq \mathrm{H}[\mu]/\epsilon] \leq \epsilon$. נגדיר אם כן את הקבוצה $S' = \{s \in S : X(s) \leq \mathrm{H}[\mu]/\epsilon\} = \{s \in S : \mu(s) \geq 2^{-\mathrm{H}[\mu]/\epsilon}\}$. לפי חסם ההסתברות מקודם (שהוא על המשלים של S') מתקיים $\Pr[S'] \geq 1-\epsilon$. לבסוף, מכיוון שמתקיים $|S'| \cdot 2^{-\mathrm{H}[\mu]/\epsilon} \leq \sum_{s\in S'} \mu(s) = \Pr[S'] \leq 1$, חייב להתקיים $|S'| \leq 2^{\mathrm{H}[\mu]/\epsilon}$ כנדרש.

שני שימושים באי־שוויון פינסקר

א. נניח שיש לנו שני משתנים מקריים X ו־Y, ונגדיר שני מרחבי הסתברות על הערכים שלהם, כפי שהוגדרו עבור הניתוח של $\mathrm{I}[X,Y]$. המרחב μ יוגדר לפי $\Pr_\mu[(\alpha,\beta)] = \Pr[X = \alpha \wedge Y = \beta]$, והמרחב ν יוגדר לפי $\Pr_\nu[(\alpha,\beta)] = \Pr[X = \alpha]\Pr[Y = \beta]$. עתה נפתח, כאשר הסכומים הם על α ו־β שיש להם סיכוי חיובי להתקבל כערך המ"מ המתאימים:

$$\begin{aligned}
\mathrm{Cov}[X,Y] = {} & \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
= {} & \sum_{\alpha,\beta} \alpha\beta\Pr[X = \alpha \wedge Y = \beta] - (\sum_{\alpha} \alpha\Pr[X = \alpha])(\sum_{\beta} \beta\Pr[X = \beta]) \\
= {} & \sum_{\alpha,\beta} \alpha\beta(\Pr_\mu[(\alpha,\beta)] - \Pr_\nu[(\alpha,\beta)]) \\
= {} & \sum_{\Pr_\mu[(\alpha,\beta)] > \Pr_\nu[(\alpha,\beta)]} \alpha\beta(\Pr_\mu[(\alpha,\beta)] - \Pr_\nu[(\alpha,\beta)]) - \sum_{\Pr_\mu[(\alpha,\beta)] < \Pr_\nu[(\alpha,\beta)]} \alpha\beta(\Pr_\nu[(\alpha,\beta)] - \Pr_\mu[(\alpha,\beta)])
\end{aligned}$$

בסוף יצא לנו הפרש של שני סכומים, כ"א מהם של איברים חיוביים. עכשיו משתמשים בנתון שערכי המ"מ הם בין 0 ל־1, ואז המחוסר חסום ע"י $\sum_{\Pr_\mu[(\alpha,\beta)] > \Pr_\nu[(\alpha,\beta)]}(\Pr_\mu[(\alpha,\beta)] - \Pr_\nu[(\alpha,\beta)]) \leq d(\mu,\nu)$, והמחסר גם חסום ע"י $\sum_{\Pr_\mu[(\alpha,\beta)] < \Pr_\nu[(\alpha,\beta)]}(\Pr_\nu[(\alpha,\beta)] - \Pr_\mu[(\alpha,\beta)]) \leq d(\mu,\nu)$. על כן $|\mathrm{Cov}[X,Y]| \leq d(\mu,\nu)$. מכאן ההמשך נובע מאי־שוויון פינסקר ומהקשר בין המידע המשותף לבין מרחק האנטרופיה היחסית המתאים:

$$|\mathrm{Cov}[X,Y]| \leq d(\mu,\nu) \leq \sqrt{\frac{1}{2}D(\mu\|\nu)} = \sqrt{\frac{1}{2}\mathrm{I}[X,Y]}$$

ב. ראשית נפתח את $\mathrm{H}[\mu]$ במושגים של $|S|$ ו־$D(\mu\|\pi)$:

$$\mathrm{H}[\mu] = -\sum_{s\in S} \mu(s)\log(\mu(s)) \;=\; -\sum_{s\in S} \mu(s)\log(\mu(s) \cdot \frac{1}{|S|} \cdot |S|)$$

$$= -\sum_{s\in S}\mu(s)\log(\frac{1}{|S|})-\sum_{s\in S}\mu(s)\log(\mu(s)\cdot|S|)=\log(|S|)-D(\mu\|\pi)$$

כל שנותר הוא להציב את אי השוויון $D(\mu\|\pi)\geq 2(d(\mu,\pi))^2$ הנובע מהעברת אגפים באי־שוויון פינסקר, וקיבלנו את המבוקש.

סאב־מודולריות אנטרופית

עבור $A,B\subseteq\{1,\ldots,n\}$, נשתמש באי־השוויון $\mathrm{H}[X|Y,Z]\leq\mathrm{H}[X|Y]$ עם ההצבות המתאימות על מנת לקבל $\mathrm{H}[X_{A\setminus B}|X_{A\cap B},X_{B\setminus A}]\leq\mathrm{H}[X_{A\setminus B}|X_{A\cap B}]$. כעת נשתמש בכלל השרשרת ונקבל:

$$\begin{aligned}\mathrm{H}[X_A]-\mathrm{H}[X_{A\cap B}] &= \mathrm{H}[X_{A\setminus B},X_{A\cap B}]-\mathrm{H}[X_{A\cap B}]=\mathrm{H}[X_{A\setminus B}|X_{A\cap B}]\\ &\geq \mathrm{H}[X_{A\setminus B}|X_{A\cap B},X_{B\setminus A}]\\ &= \mathrm{H}[X_{A\setminus B},X_{A\cap B},X_{B\setminus A}]-\mathrm{H}[X_{A\cap B},X_{B\setminus A}]=\mathrm{H}[X_{A\cup B}]-\mathrm{H}[X_B]\end{aligned}$$

העברת אגפים תתן עתה את המבוקש.

רעש רקע

עבור אי השוויון הראשון, נשים לב שמספיק להוכיח שלכל β שעבורו $\Pr[Z=\beta]>0$ מתקיים אי השוויון $\mathrm{H}[X+Y|Z=\beta]\geq\mathrm{H}[Y|Z=\beta]$. נשים לב שגם בהתניה של מרחב ההסתברות על $Z=\beta$ מתקיים ש־X הוא ב"ת ב־Y (כי X היה ב"ת בצירוף של Y ו־Z). על כן, מספיק להראות שלכל זוג משתנים ב"ת X',Y' מתקיים $\mathrm{H}[X'+Y']\geq\mathrm{H}[Y']$ (אצלנו אלו יהיו ההתניות של X ו־Y על $Z=\beta$).

לשם כך, נשים לב שמתקיים $\mathrm{H}[X',Y']=\mathrm{H}[X']+\mathrm{H}[Y']$ כי אלו מ"מ ב"ת, וכן $\mathrm{H}[X',Y']=\mathrm{H}[X',X'+Y']$, וזאת מכיוון שהפונקציה $f(a,b)=(a,a+b)$ היא הפיכה (לפי הפונקציה $f^{-1}(a,b)=(a,b-a)$). על מנת להבין את השוויון עבור פונקציות חח"ע של מ"מ אפשר או להיזכר בהגדרה המקורית של אנטרופיה (שלא תלויה בערכים של המ"מ עצמם אלא רק בהסתברות לכל ערך), או להשתמש באי השוויון מהשיעור עבור פונקציות של מ"מ לקבלת $\mathrm{H}[X',Y']\geq\mathrm{H}[f(X',Y')]=\mathrm{H}[X',X'+Y']$ וגם $\mathrm{H}[X',X'+Y']\geq\mathrm{H}[f^{-1}(X',X'+Y')]=\mathrm{H}[X',Y']$ להוכחת השוויון.

מהשוויונים הנ"ל מתקיים $\mathrm{H}[X']+\mathrm{H}[Y']=\mathrm{H}[X',X'+Y']\leq\mathrm{H}[X']+\mathrm{H}[X'+Y']$ לפי תת־חיבוריות של שרשור מ"מ, ומהעברת אגפים מקבלים $\mathrm{H}[X'+Y']\geq\mathrm{H}[Y']$.

עבור אי־השוויון השני נשים לב שלכל α שעבורו $\Pr[X=\alpha]>0$ מתקיים $\mathrm{H}[Y|Z,X=\alpha]=\mathrm{H}[Y|Z]$, בגלל אי־התלות של X בצירוף של Y ו־Z. על כן $\mathrm{H}[Y|Z]=\mathrm{E}_{\alpha\sim X}\mathrm{H}[Y|Z,X=\alpha]=\mathrm{H}[Y|Z,X]$. כמו כן מתקיים $\mathrm{H}[Z,X]=\mathrm{H}[Z,X+Z]$ (שוב לפי כך ש־$f(a,b)=(a,a+b)$ היא חח"ע), וגם $\mathrm{H}[Y,Z,X]=\mathrm{H}[Y,Z,X+Z]$ (שוב לפי פונקציה חח"ע), ומשני אלו באמצעות שימוש בכלל השרשרת של אנטרופיה מותנית נקבל את השוויון $\mathrm{H}[Y|Z,X]=\mathrm{H}[Y,Z,X]-\mathrm{H}[Z,X]=\mathrm{H}[Y,Z,Z+X]-\mathrm{H}[Z,Z+X]=\mathrm{H}[Y|Z,Z+X]$.

מכל אלו קיבלנו סה"כ $\mathrm{H}[Y|Z]=\mathrm{H}[Y|Z,Z+X]$. לבסוף, לפי אי־השוויון מהשיעור על התניות על שרשורים של מ"מ, נקבל $\mathrm{H}[Y|Z,Z+X]\leq\mathrm{H}[Y|Z+X]$, להשלמת ההוכחה.

לא תרבו

ראשית, נסמן ב־$B = \{i : \Pr_\mu[i] \geq 1/\alpha n\}$ את קבוצת האינדקסים המתקבלים בהסתברות גבוהה מ־$1/\alpha n$. נשים לב קודם כל שמתקיים $|B| \leq \alpha n$, פשוט כי סכום ההסתברויות לא יכול להיות גדול מ־1. על כן לפי נתוני השאלה יש מילה $w \in C$ שמתאפסת על B. מכאן שבפרט $\Pr_{i\sim\mu}[i \in B] \leq \frac{1}{10}$, כי אחרת לא יכול להיות שיתקיים $\Pr_\mu[w_i = 1] \geq \frac{9}{10}$ כפי שנתון.

עתה נסתכל על התהליך הבא: מגרילים את $w \in C$ באופן יוניפורמי מהקבוצה הנ"ל, ובאופן ב"ת בהגרלת w מגרילים את i לפי μ. נשים לב שמתקיים $\mathrm{H}[w] = \log|C|$ לפי הידוע על אנטרופיה של התפלגות יוניפורמית, וכן שמתקיים $\Pr[w_i = 1] \geq \frac{9}{10}$. אי השוויון השני נכון ל־w המוגרל אקראית ע"י "מיצוע" של אי השוויון שמתקיים לכל $w \in C$ קבוע (בעצם זו נוסחת ההסתברות השלמה). על מנת לסיים נרצה לחסום את $\mathrm{H}[w]$, ודרכו את $|C|$.

נסמן ב־ν את המרחב של הגרלה יוניפורמית של $w \in C$, ונסמן ב־$J = \{j : \Pr_{w\sim\nu}[w_j = 1] \geq \frac{7}{10}\}$ את קבוצת האינדכסים שעבורם ההגרלה של w תתן ערך 1 בהסתברות $\frac{7}{10}$ לפחות. מתקיים $\Pr_{i\sim\mu}[i \in J] \geq \frac{2}{10}$, כי המאורע "$w_i = 1$" במרחב המשולב של הגרלת w ו־i מוכל באיחוד המאורע "$i \in J$" עם המאורע "$w_i = 1 \wedge i \notin J$", ולכן מתקיים $\frac{9}{10} \leq \Pr[w_i = 1] \leq \Pr[i \in J] + \Pr[w_i = 1 \wedge i \notin J] \leq \Pr[i \in J] + \frac{7}{10}$.

בפרט, ע"פ הכלל על איחוד מאורעות, חייב להתקיים $\Pr_{i\sim\mu}[i \in J \setminus B] \geq \frac{1}{10}$. מכיוון שבקבוצה הנ"ל אין איברים עם הסתברות גבוהה מ־$1/\alpha n$ (לפי הגדרת B), מתקיים $|J| \geq |J \setminus B| \geq \alpha n/10$. לפי סאב־אדיטיביות של אנטרופיה אנחנו יודעים שמתקיים $\mathrm{H}[w] = \mathrm{H}[w_1, \ldots, w_n] \leq \sum_{i=1}^n \mathrm{H}[w_i]$. לבסוף, נשים לב שעבור כל i מתקיים $\mathrm{H}[w_i] \leq 1$, בעוד שעבור $i \in J$ מתקיים $\mathrm{H}[w_i] \leq H(\frac{7}{10}) < 1$ (הביטוי באמצע הוא פונקצית האנטרופיה המספרית שהוגדרה בשיעור). נסמן $\beta = \frac{\alpha}{10}(1 - H(\frac{7}{10})) > 0$, ונקבל מאלו $\log|C| = \mathrm{H}[w] \leq \sum_{i=1}^n \mathrm{H}[w_i] \leq (1-\beta)n$. לקיחת חזקה של שני האגפים תתן $|C| \leq 2^{(1-\beta)n}$ כמבוקש.

דומה בדומה

נראה כיוון אחד, $\mathrm{H}[\mu] - \mathrm{H}[\nu] \leq d(\mu,\nu)\log(n-1) + H(d(\mu,\nu))$. החסם מהצד השני סימטרי לחלוטין.

נשתמש בשאלה "התפלגויות מותנות" (מהפרק על בניה וניתוח של מרחבי הסתברות) ונבנה מרחב הסתברות עם שני משתנים מקריים X ו־Y, כאשר X מתפלג כמו μ, Y מתפלג כמו ν, ומתקיים $\Pr[X \neq Y] = d(\mu,\nu)$. עבור אלו מתקיים $\mathrm{H}[\mu] - \mathrm{H}[\nu] = \mathrm{H}[X] - \mathrm{H}[Y] \leq \mathrm{H}[X,Y] - \mathrm{H}[Y] = \mathrm{H}[X|Y]$. כל שנותר הוא להשתמש באי־שוויון פאנו (בגרסה ה"מקוצרת"):

$$\mathrm{H}[X|Y] \leq \Pr[X \neq Y]\log(n-1) + H(\Pr[X \neq Y]) = d(\mu,\nu)\log(n-1) + H(d(\mu,\nu))$$

סודות ושקרים

כפי שנאמר בהדרכה, התשובה הסופית של האלגוריתם היא פונקציה של סדרת התשובות שניתנו ב־q השלבים. נסמן את התשובה בשלב ה־i ב־$a_i \in \{0,1\}$ (0 בשביל "לא" ו־1 בשביל "כן"), ואת סידרת כל התשובות ב־$X = (a_1, \ldots, a_q) \in \{0,1\}^q$. נניח ש־$k$ נבחר באופן יוניפורמי מתוך $\{1, \ldots, n\}$, ונסמן אותו כמשתנה מקרי $Y = k$. לבסוף נסמן את התשובה של האלגוריתם עצמו גם כמשתנה מקרי, $Z = k'$. שימו לב בפרט ש־Z היא פונקציה של X. נניח שהאלגוריתם מקיים $Y = Z$ בהסתברות לפחות $1-\alpha$ (את הערך הספציפי של α נבחר לקראת סוף ההוכחה), ונראה חסם תחתון על האנטרופיה $\mathrm{H}[X]$. מכיוון שתמיד מתקיים $\mathrm{H}[X] \leq \log|\{0,1\}^q| = q$, זה ייתן לנו חסם תחתון על q.

ראשית נחסום את $\mathrm{H}[X,Y] - \mathrm{H}[X] = \mathrm{H}[Y|X] = \mathrm{H}[Y|X,Z] \leq \mathrm{H}[Y|Z] = \mathrm{H}[Y,Z] - \mathrm{H}[Z]$ (השתמשנו בשוויון השני בכך ש־Z הוא פונקציה של X): מכיוון ש־Z שווה ל־Y בהסתברות לפחות $1-\alpha$, לכל k מתקיים

$\Pr[Y=Z=k] \geq (1-\alpha)/n$. סה"כ ההסתברות לכל המאורעות מהצורה $Y=k \neq k'=Z$ חסום ע"י α. נשים לב שתחת האילוצים האלו, הביטוי

$$\mathrm{H}[Y,Z] = \sum_{k=1}^{n} \Pr[Y=Z=k] \log \frac{1}{\Pr[Y=Z=k]} + \sum_{k \neq k'} \Pr[Y=k \wedge Z=k'] \log \frac{1}{\Pr[Y=k \wedge Z=k']}$$

יקבל את המקסימום כאשר ההתפלגות על ערכי Y,Z היא יוניפורמית בהתניה על המקרה $Y \neq Z$ (בדומה לכך שהאנטרופיה של מ"מ כל שהוא מקבלת את המקסימום כאשר הוא מתפלג יוניפורמית). זה נותן לנו

$$\mathrm{H}[Y,Z] \leq \log(n) + n(n-1) \cdot \frac{\alpha}{n(n-1)} \log \frac{n(n-1)}{\alpha} \leq \log(n) + \alpha(\log(1/\alpha) + 2\log(n))$$

נניח עתה ש־n גדול מספיק (ביחס ל־α שאח"כ נבחר) על מנת לקיים $\log(1/\alpha) < \log(n)$. לכל k מתקיים $\Pr[Z=k] \geq \Pr[Y=Z=k] \geq (1-\alpha)/n$, ואם נניח גם ש־$n \geq 3$ נובע מזה (ומהמונוטוניות של $x \log \frac{1}{x}$ בתחום $0 < x < \frac{1}{e}$) שמתקיים $\Pr[Z=k] \log \frac{1}{\Pr[Z=k]} \geq \frac{1-\alpha}{n} \log \frac{n}{1-\alpha}$. כשסוכמים על $1 \leq k \leq n$ מקבלים $\mathrm{H}[Z] \geq (1-\alpha) \log \frac{n}{1-\alpha} > (1-\alpha)\log(n)$, וסה"כ $\mathrm{H}[X,Y] - \mathrm{H}[X] \leq \mathrm{H}[Y,Z] - \mathrm{H}[Z] \leq 4\alpha \log(n)$.

עתה נחסום מהצד השני את $\mathrm{H}[X,Y] = \mathrm{H}[X|Y] + \mathrm{H}[Y]$. לפי ההגדרות מתקיים $\mathrm{H}[Y] = \log(n)$. כמו כן, לפי הנתון שההסתברות ל"שקר" כל פעם שהאלגוריתם שואל שאלה היא בדיוק $\frac{1}{10}$ (באופן ב"ת בשאלות קודמות), אפילו בהתניה על ערך ספציפי של Y מתקיים $\mathrm{H}[X|Y=k] = qH(\frac{1}{10})$ (בצד ימין יש את "פונקצית האנטרופיה" המספרית, ובפרט יש שם מקדם קבוע גדול מ־0), ולכן $\mathrm{H}[X|Y] = qH(\frac{1}{10})$.

סה"כ אנחנו מקבלים $\mathrm{H}[X,Y] = qH(\frac{1}{10}) + \log(n)$. יחד עם חסם ההפרש בין $\mathrm{H}[X]$ ו־$\mathrm{H}[X,Y]$ נקבל $q \geq \mathrm{H}[X] \geq \mathrm{H}[X,Y] - 4\alpha\log(n) = qH(\frac{1}{10}) + (1-4\alpha)\log(n)$, ובהעברת אגפים $q \geq \frac{1-4\alpha}{1-H(1/10)} \log(n)$. עבור $\alpha = H(\frac{1}{10})/8$ ו־$\beta = H(\frac{1}{10})/16$ למשל נקבל את הנדרש $q \geq (1+\beta)\log(n)$.

הערה: חישוב יותר מדוייק של מספר השאלות הקטן ביותר האפשרי (חסם עליון ותחתון), כאשר ההסתברות לשקר היא p, הושג ע"י Rényi בשנת 1961, ועומד על $(1 \pm o(1))\log(n)/(1-H(p))$ עבור שגיאה חסומה ע"י קבוע.

גן השבילים המתפצלים

נגדיר את $X_0,\ldots,X_k$ להיות הילוך מקרי ללא חזרות לאחור, בדיוק כמו בשאלה "לא חוזרים לאחור" מפרק השאלות על הילוכים מקריים. מהשאלה הזו נובע שההתפלגות הלא מותנה על X_k נתונה לפי $\Pr[X_k = v] = \frac{d(v)}{2|E|}$. עתה נבדוק מהי האנטרופיה המותנית $\mathrm{H}[X_1,\ldots,X_k|X_0]$.

מצד אחד, בגלל שאין מעגלים מגודל קטן מ־$2k+1$, אם אנחנו יודעים את X_0 ואת X_k, אז אנחנו בהכרח יודעים גם את $X_1,\ldots,X_k$ (עבור זוג X_0,X_k לגיטימי יהיה מסלול יחיד מאורך k ביניהם). על כן מתקיים $\mathrm{H}[X_1,\ldots,X_k|X_0] = \mathrm{H}[X_k|X_0] \leq \mathrm{H}[X_k] \leq \log(n)$.

מצד שני, נשתמש $k-1$ פעמים בכלל השרשרת לקבלת $\mathrm{H}[X_1,\ldots,X_k|X_0] = \sum_{i=1}^{k} \mathrm{H}[X_i|X_0,\ldots,X_{i-1}]$. כמו כן, עבור $i \geq 2$, מכיוון שאופן ההגרלה של X_i תלוי אך ורק בערכים של X_{i-1} ו־X_{i-2}, מתקיים $\mathrm{H}[X_i|X_0,\ldots,X_{i-1}] = \mathrm{H}[X_i|X_{i-2},X_{i-1}]$. הוכחה מהירה של השוויון הזה:

$$\mathrm{H}[X_i|X_0,\ldots,X_{i-1}] = \sum_{\alpha_0,\ldots,\alpha_{i-1}} \mathrm{H}[X_i|X_0=\alpha_0,\ldots,X_{i-1}=\alpha_{i-1}]\Pr[X_0=\alpha_0,\ldots,X_{i-1}=\alpha_{i-1}]$$

$$\begin{aligned}
&= \sum_{\alpha_0,\ldots,\alpha_{i-1}} \mathrm{H}[X_i|X_{i-2}=\alpha_{i-2},X_{i-1}=\alpha_{i-1}]\Pr[X_0=\alpha_0,\ldots,X_{i-1}=\alpha_{i-1}] \\
&= \sum_{\alpha_{i-2},\alpha_{i-1}} \mathrm{H}[X_i|X_{i-2}=\alpha_{i-2},X_{i-1}=\alpha_{i-1}]\Pr[X_{i-2}=\alpha_{i-2},X_{i-1}=\alpha_{i-1}] \\
&= \mathrm{H}[X_i|X_{i-2},X_{i-1}]
\end{aligned}$$

עתה נחסום את המחוברים לפי ההגדרה של אנטרופיה מותנית כממוצע של אנטרופיות מותנות על ערכים ספציפיים.

מחובר ראשון: $\mathrm{H}[X_1|X_0]=\sum_{v\in V}\log(d(v))\cdot\frac{d(v)}{2|E|}=\frac{n}{2|E|}\cdot\frac{1}{n}\sum_{v\in V}d(v)\log(d(v))$, ובהפעלת אי־שוויון ינסן על הפונקציה $f(z)=z\log(z)$ מקבלים $\mathrm{H}[X_1|X_0]\geq\frac{n}{2|E|}d\log(d)=\log(d)$, כאשר כזכור $d=\frac{1}{n}\sum_{v\in V}d(v)$, "תוחלת הדרגה" מעל בחירה יוניפורמית של צומת.

את המחובר $\mathrm{H}[X_i|X_{i-2},X_{i-1}]$ כאשר $1<i\leq k$, מפתחים כשהפעם אי שוויון ינסן מופעל על הפונקציה $f(z)=z\log(z-1)$ (שימו לב שעבור $z\geq 2$ הנגזרת השניה של הפונקציה היא $\frac{1}{z-1}-\frac{1}{(z-1)^2}\geq 0$). זוגות הערכים האפשריים עבור (X_{i-2},X_{i-1}) הם כל הקשתות האפשריות של הגרף (כל אחת בשני הסדרים האפשריים), וכל זוג כזה מתקבל בהסתברות $\frac{1}{2|E|}$ בדיוק (ראו את השאלה "לא חוזרים לאחור"). החסם על $\mathrm{H}[X_i|X_{i-2},X_{i-1}]$ הוא $\sum_{uv\in E}\log(d(v)-1)\cdot\frac{1}{2|E|}=\sum_{v\in V}\log(d(v)-1)\cdot\frac{d(v)}{2|E|}\geq\frac{n}{2|E|}d\log(d-1)=\log(d-1)$.

סה"כ קיבלנו $\mathrm{H}[X_1,\ldots,X_k|X_0]\geq\log(d)+(k-1)\log(d-1)$. יחד עם אי השוויון הראשון (החסם מלמעלה) נקבל $\log(d)+(k-1)\log(d-1)\leq\log(n)$, והעלאת חזקה בשני האגפים תתן לנו $d(d-1)^{k-1}\leq n$ כנדרש.

משפחות נחתכות במשולשים

ראשית נתאים את אי השוויון מהתרגיל "הטלות נחתכות" לצרכינו. נקבע כי $S_i=\{0,1\}$ כולן, ואז ניתן לזהות את הוקטורים ב־$\mathcal{F}$ עם תתי קבוצות של $[n]$. כעת מקבלים שאם $\mathcal{G}=\{G_1,\ldots,G_m\}$ אוסף של תתי קבוצות של $[n]$ כך שכל איבר ב־$[n]$ מופיע בלפחות k קבוצות ב $\mathcal{G}$, ואם נסמן $\mathcal{F}_i=\{F\cap G_i : F\in\mathcal{F}\}$, אז מתקיים $|\mathcal{F}|^k\leq\prod_{i=1}^m|\mathcal{F}_i|$.

כעת נוכל להוכיח את הטענה. נסמן ב־N את קבוצת כל $\binom{t}{2}$ הזוגות הלא־סדורים של איברים ב־$[t]$, ונביט ב־$\mathcal{F}$ כמשפחה של תתי קבוצות של N. תהא $\mathcal{G}$ משפחת כל תתי הקבוצות של N שהן קבוצת קשתות של איחוד זר של שני גרפים מלאים, אחד על $\lfloor t/2\rfloor$ צמתים והשני על $\lceil t/2\rceil$ צמתים. נסמן ב־ $s=\binom{\lceil t/2\rceil}{2}+\binom{\lfloor t/2\rfloor}{2}$ את מספר הקשתות בגרף כזה. נסמן גם $|\mathcal{G}|=m$. משיקולי סימטריה, כל קשת ב־N נמצאת בבדיוק $k=sm/\binom{t}{2}$ גרפים מ־$\mathcal{G}$.

כעת, נשים לב שלכל $G\in\mathcal{G}$ מתקיים שהגרף המשלים לו הוא חסר משולשים, ולכן מכיוון שכל שני גרפים ב־$\mathcal{F}$ נחתכים במשולש, הם גם חייבים לחלוק קשת עם G, וטיעון זה יפה לכל $G\in\mathcal{G}$. כך, אם נחזור לסימונים של תחילת ההוכחה, הגודל של כל $\mathcal{F}_i$ הוא לכל היותר 2^{s-1}, שכן מכיוון שלכל שני גרפים ב־$\mathcal{F}$ יש לפחות קשת אחת משותפת שנמצאת גם ב־G_i, יכולים להיות לכל היותר 2^{s-1} גרפים ב־$\mathcal{F}$ הנבדלים על קשתות G_i (שכן לא יכולה להיות ב־$\mathcal{F}_i$ קבוצה ומשלימתה, שהרי כאלה קבוצות היו מתאימות לשני גרפים ב־$\mathcal{F}$ שאינם נחתכים על G_i כלל). נציב זאת באי השוויון מההתחלה ונקבל $|\mathcal{F}|^{sm/\binom{t}{2}}\leq\left(2^{s-1}\right)^m$, ובלקיחת שורשים משני הצדדים $|\mathcal{F}|\leq 2^{\binom{t}{2}-\binom{t}{2}/s}$. לפי משולש פסקל $s<\frac{1}{2}\binom{t}{2}$ וכך $|\mathcal{F}|<2^{\binom{t}{2}-2}=\frac{1}{4}2^{\binom{t}{2}}$.

משפחות נחתכות בשידוכים

נעקוב אחרי פתרון התרגיל "משפחות נחתכות במשולשים". נסמן ב־N את קבוצת כל $\binom{2n}{2}$ הזוגות הלא סדורים של איברים ב־$\{1,2,\ldots,2n\}$. נביט ב־$\mathcal{F}$ כמשפחה של תתי קבוצות של N. תהא $\mathcal{G}$ משפחת כל תתי הקבוצות של

N המתאימות לכוכבים (פורשים) על $2n$ צמתים. מספר הקשתות בגרף כזה הוא $s = 2n-1$, ונסמן גם $|\mathcal{G}| = m$. שוב, משיקולי סימטריה כל קשת ב־N נמצאת בבדיוק $sm/\binom{2n}{2}$ גרפים ב־$\mathcal{G}$.

נשים לב שלכל $G \in \mathcal{G}$ מתקיים שהשידוך המקסימלי במשלים שלו הוא עם $n-1$ קשתות (לא ניתן להשתמש במרכז הכוכב) ולכן כל שני גרפים ב־$\mathcal{F}$ נחתכים בקשת מ־G. לכן ניתן לחזור על אותם טיעונים ולקבל ש־$|\mathcal{F}| \le 2^{\binom{2n}{2}-\binom{2n}{2}/s} = 2^{\binom{2n}{2}-n}$.

הילוכים מקריים

הילוך מהוסס על הקוביה

נגדיר הילוך מקרי על הקוביה הבוליאנית $\{0,1\}^n$ באופן הבא: X_0 הוא הווקטור $(0,\ldots,0)$ בהסתברות 1. בהינתן X_i, אנו נבחר בהסתברות $\frac{1}{2}$ את X_{i+1} להיות זהה לו, ובהסתברות $\frac{1}{2}$ נגדיר את X_{i+1} ע"י כך שנבחר באופן מקרי ויוניפורמי קורדינטה של X_i ונהפוך את ערכה, כששאר הקורדינטות ישארו אותו דבר. הוכיחו שעבור כל $\epsilon > 0$ קבוע המרחק בין ההתפלגות של X_t לבין ההתפלגות היוניפורמית על $\{0,1\}^n$ קטן מ־ϵ, עבור $t = O(n\log n)$.

שתי שאלות על הילוכים מקריים מהוססים

נניח שאנו מבצעים על גרף G (במקום הילוך מקרי רגיל) את הפרוצדורה הבאה: בכל שלב, בסיכוי חצי נבצע את ההילוך המקרי לפי בחירה של שכן מקרי של הצומת הנוכחי, ובסיכוי חצי פשוט נישאר צעד אחד נוסף בצומת הנוכחי. נניח גם שהגרף עליו מתבצע ההילוך הוא קשיר. הוכיחו עבור הילוך כזה שני דברים.

- הילוך כזה תמיד יתכנס להתפלגות הסטציונרית (אותה אחת כמו עבור הילוך מקרי רגיל), גם אם הגרף הוא 2־צביע.
- זמני הביקור (הממוצעים) בהילוך כזה הם בדיוק כפולים מאלו של הילוך מקרי רגיל.

לא חוזרים לאחור

בהינתן גרף $G = (V,E)$ שכל הצמתים שלו מדרגה 2 לפחות, נגדיר את ההילוך הלא־חזרתי על G בצורה הבאה: $X_0 \in V$ נבחר לפי ההפלגות הסטציונרית, ז"א $\Pr[X_0 = v] = d(v)/2|E|$. אחר כך $X_1 \in V$ נבחר יוניפורמית מתוך $d(X_0)$ השכנים של הצומת שנבחר עבור X_0. בשלב ה־i עבור $i > 1$, אחרי שכבר בחרנו את $X_0,\ldots,X_{i-1}$, הערך של $X_i \in V$ נבחר באופן יוניפורמי מתוך השכנים של X_{i-1} **פרט** ל־X_{i-2} (סה"כ $d(X_{i-1})-1$ אפשרויות).

הראו שלכל i ההתפלגות הלא־מותנה של X_i היא ההתפלגות הסטציונרית, ז"א $\Pr[X_i = v] = d(v)/2|E|$.

קשיים בהתקדמות

נגדיר הילוך מקרי מוטה על הישר. נקבע $X_0 = 0$, ולכל $i > 0$ בהסתברות $\frac{1}{3}$ נקבע $X_i = X_{i-1}+1$ ובהסתברות $\frac{2}{3}$ נקבע $X_i = \max\{0, X_{i-1}-1\}$, ללא תלות בבחירות הקודמות. נסמן ב־t_k את תוחלת ההפרש בין ה־i הקטן ביותר כך ש־$X_i = k-1$ וה־j הקטן ביותר עבורו $X_j = k$. לדוגמא:

$$t_1 = \sum_{r=1}^{\infty}\frac{1}{3}(\frac{2}{3})^{r-1}r = \sum_{r=1}^{\infty}\sum_{s=1}^{r}\frac{1}{3}(\frac{2}{3})^{r-1} = \sum_{s=1}^{\infty}\sum_{r=s}^{\infty}\frac{1}{3}(\frac{2}{3})^{r-1} = \sum_{s=1}^{\infty}(\frac{2}{3})^{s-1} = 3$$

חשבו את t_k.

טיול בגרף נאה

נתון ש־s ו־t הם צמתים בגרף 2־קשיר (בצמתים) ו־3־רגולרי (דרגות כל צמתיו 3) בעל n צמתים. הראו שמתקיים $k_{st} \le \frac{3}{4}n^2$. ניתן להסתמך על ידע בפיזיקה.

בלי הרבה נפנוף ידיים

נניח ש־$X_0, X_1, \ldots$ הוא הילוך מקרי על הגרף (הלא מכוון והקשיר) G, אשר יוצא מ־v (ז"א $X_0 = v$ בהסתברות 1). נסמן ב־T את המ"מ המקבל את זמן החזרה הראשון ל־v לאחר היציאה ממנו, ז"א $T = \min\{t : X_t = v \wedge t > 0\}$, ונסמן $\tau = \mathrm{E}[T]$. לאחר פתרון תרגיל זה תדעו איך מוכיחים שמתקיים $\tau = 1/\pi_v$ (כאשר כאן π מסמן את ההתפלגות הסטציונרית).

הערות: בשאלות יש סימונים מהצורה "$o(1)$", אולם כדאי להעביר את הפורמליזם לאחד מהצורה "לכל $\epsilon > 0$ קבוע מתקיים עבור s גדול דיו... (שיכול להיות תלוי ב־G)".

עבור s גדול דיו, נסמן ב־H_s את מספר הפעמים שנכנסנו ל־v ב־s הצעדים הראשונים של ההילוך המקרי שלנו, ז"א $H_s = |\{i : X_i = v \wedge 1 \le i \le s\}|$.

- הראו שמתקיים $\mathrm{E}[H_s] = (1 \pm o(1))s\pi_v$.
- הראו שבסיכוי $1 - o(1)$, מתקיים $H_s = (1 \pm o(1))s/\tau$.

מהסעיף השני נובע $\mathrm{E}[H_s] = (1 \pm o(1))s/\tau$ מכיוון שבכל מקרה H_s מקבל ערכים בין 0 ל־s, ואז משני הסעיפים יחד נובע המבוקש.

רמז: עבור הסעיף השני, לכל $j \ge 1$ אפשר לבדוק את תוחלת מספר הצעדים בין הביקור ה־$j - 1$ וה־j ב־v.

גם זו הרמונית

נתון גרף G (לא מכוון וקשיר) עם קבוצת צמתים V. לכל זוג צמתים $s \ne t$ וצומת v נגדיר את $U_{st}(v)$ להיות תוחלת מספר המעברים ב־v המבוצע ע"י הילוך מקרי המתחיל ב־s, עד לפגיעה הראשונה ב־t. אנו סופרים את היציאה מ־s (ז"א שמתקיים $U_{st}(s) \ge 1$) אולם לא את הכניסה ל־t (כך ש־$U_{st}(t) = 0$). נגדיר עתה את $\phi_{st} : V \to \mathbb{R}$ לפי $\phi_{st}(v) = U_{st}(v)/d(v)$. הראו שזוהי פונקציה הרמונית עם שפה $\{s, t\}$.

הערה: הפונקציה U_{st} משמשת בהוכחה המקורית של Tetali עבור h_{st}. שימו לב שמתקיים $h_{st} = \sum_{v \in V} U_{st}(v)$.

פתרונות לתרגילים על הילוכים מקריים

הילוך מהוסס על הקוביה

ההוכחה נעשית בשיטת הצימוד. ראשית נשים לב שניתן היה להגדיר באופן שקול את ההילוך המקרי $X_0, X_1, \ldots$ ע"י כך שבשלב ה־i קודם נבחר באופן יוניפורמי קואורדינטה $1 \le j_i \le n$, ורק אחר כך נחליט בהסתברות $\frac{1}{2}$ אם להפוך את הקורדינטה ה־j_i. עתה נגדיר הילוך מקרי שני $Y_0, Y_1, \ldots$ על הקוביה $\{0,1\}^n$ באופן הבא: Y_0 יבחר באופן מקרי ויוניפורמי מהקוביה. בשלב ה־i נבדוק מהי הקורדינטה שנבחרה כשבחרנו את X_{i+1} לפי X_i. אם X_i ו־Y_i זהים על הקורדינטה הנ"ל אז לבחירת Y_{i+1} אנו נהפוך את הקורדינטה אם ורק אם $X_{i+1} \ne X_i$ (ז"א שבסיכוי $\frac{1}{2}$ נהפוך אותה בשני ההילוכים, ובסיכוי $\frac{1}{2}$ לא נהפוך אותה באף אחד מהם). אם X_i ו־Y_i שונים על הקורדינטה הנ"ל אז נהפוך אותה עבור Y_{i+1} אם ורק אם $X_{i+1} = X_i$.

ניתן לראות בשלב זה ש־$Y_0, Y_1, \ldots$ לכשעצמו הוא הילוך מקרי עם אותה מטריצת מעבר כמו $X_0, X_1, \ldots$. מכיוון שהתפלגות של Y_0 היא ההתפלגות הסטציונרית של ההילוך, ההתפלגות (הלא מותנה) של Y_t תהיה זהה לה לכל $0 \le t$. נראה עתה שהמאורע $X_t = Y_t$ יקרה בהסתברות העולה על $1 - \epsilon$ עבור $t = n \ln(n/\epsilon) = O(n \log n)$ לסיום ההוכחה.

לכל $1 \le j \le n$ נגדיר את המאורע A_j כמאורע שהקורדינטה j נבחרה להפיכה אפשרית במעבר מ־X_i ל־X_{i+1} עבור $0 \le i < t$ כל שהוא. ניתן לראות שאם A_j מתקיים אז X_t ו־Y_t יהיו שווים על הקורדינטה ה־j, ולכן אם $\bigwedge_{j=1}^n A_j$ מתקיים אז בפרט $X_t = Y_t$. אולם $\Pr[\neg A_j] = (1 - \frac{1}{n})^t < \epsilon/n$, ולכן $\Pr[\bigwedge_{j=1}^n A_j] > 1 - \epsilon$ כנדרש.

שתי שאלות על הילוכים מקריים מהוססים

התכנסות להילוך הסטציונרי

ישנן שתי שיטות להוכיח זאת. נתחיל מהשיטה האלגברית: נניח ש־P היא מטריצת המעבר של ההילוך המקרי (הלא־מהוסס) על הגרף ששלנו, ו־P' היא מטריצת המעבר של ההילוך המהוסס. מתקיים אם כן $P' = \frac{1}{2}(P+I)$ כאשר I מסמנת את מטריצת היחידה. מכאן שיש ל־P' אותם ווקטורים עצמיים כמו ל־P, כאשר עבור כל ערך עצמי λ של P יהיה ערך עצמי $\lambda' = \frac{1}{2}(\lambda+1)$ של P'.

כל שנותר עתה הוא להיזכר בכך שכל הערכים העצמיים של P הם ממשיים (הוכח בכיתה), שערכם הוא בין 1 ל־-1 (הוכח בתרגול, זה נובע מכך שהמדובר במטריצה שסכומי השורות שלה הם 1), ושיש רק ווקטור עצמי אחד, זה של ההתפלגות הסטציונרית, שעבורו הע"ע הוא 1 (זה נובע מקשירות הגרף, עובדה זו הוכחה בתרגול). נשים לב עתה שעבור $-1 \le \lambda < 1$ מתקבל $|\lambda'| < 1$ ורק עבור $\lambda = 1$ מתקבל $\lambda' = 1$, ומכאן של־P' כל הערכים העצמיים קטנים מ־1 בערך מוחלט, פרט להתפלגות הסטציונרית, שהיא הווקטור העצמי היחידי שערכו העצמי הוא 1. מכאן נובע המבוקש.

עתה נסקור בקצרה את שיטת ההוכחה השניה. אפשר להשתמש בשיטת הצימוד: אנו נגדיר שני הילוכים מקריים מהוססים $\underline{X} = X_0, X_1, \ldots$ ו־$\underline{Y} = Y_0, Y_1, \ldots$ אשר לשניהם יהיו את מטריצת המעבר P', ההתפלגות הבלתי־מותנה של Y_0 (ולכן של כל Y_i) היא ההתפלגות הסטציונרית, ההתפלגות של $\underline{X}$ היא זו של ההילוך המקרי המקורי (מושג ע"י קביעת ההתפלגות של X_0 להיות זו של ההילוך המקורי באופן ב"ת ב־Y_0), וכן מתקיים $\Pr[Y_i = X_i] \to 1$.

לשם כך נבחר את (X_0, Y_0) כמתואר למעלה, ועתה נתאר את המעבר (המקרי) מ־(X_i, Y_i) ל־(X_{i+1}, Y_{i+1}). אם $X_i \neq Y_i$, אז בסיכוי $\frac{1}{2}$ נקבע את $X_{i+1} = X_i$ ואת Y_{i+1} להיות שכן מקרי של Y_i לפי הגרף שלנו, ובסיכוי $\frac{1}{2}$ נקבע את $Y_{i+1} = Y_i$ ואת X_{i+1} להיות שכן מקרי של X_i לפי הגרף. אם $X_i = Y_i$, אז בסיכוי $\frac{1}{2}$ נקבע $X_{i+1} = Y_{i+1} = X_i$, ובסיכוי $\frac{1}{2}$ נקבע את $X_{i+1} = Y_{i+1}$ להיות שכן מקרי של X_i. ההוכחה ששאר התכונות מתקיימות (מטריצות המעבר הלא מותנות של $\underline{X}$ ושל $\underline{Y}$, וכן התכנסות ההסתברות עבור המאורע $X_i = Y_i$) מושארת כתרגיל לקורא (עבור ההתכנסות, שימו לב שתמיד ניתן לחסום מלמטה את הסיכוי ש־$X_{i+n} = Y_{i+n}$ לכל צמד ערכים אפשרי של $X_i \neq Y_i$; לחילופין, אפשר לתאר את הצמדים (X_i, Y_i) לפני המפגש כהילוך מקרי רגיל על גרף מתאים, עם ריבוע מספר הצמתים של G, ולחסום ע"י זמני פגיעה).

זמני הביקור

ניתן לעשות רדוקציה של השאלה עבור הילוך מקרי מהוסס על הגרף G לשאלה על הילוך מקרי רגיל על גרף חדש G'. הגרף G' יוגדר ע"י כך שנהפוך כל קשת $e = v_i v_j$ של G למסלול מאורך 2 שבמרכזו צומת חדש u_e. אם ל־G היו n צמתים ו־m קשתות, אז לגרף החדש יהיו $n' = n + m$ צמתים ו־$m' = 2m$ קשתות. בנוסף, שימו לב שאם בגרף הישן ההתנגדות השקולה בין s ל־t (כפי שהוגדרה בהרצאה) היא R_{st}, אז בגרף החדש זו תהיה $R'_{st} = 2R_{st}$.

עתה נסמן ב־$\underline{Y} = Y_0, Y_1, \ldots$ הילוך מקרי (לא מהוסס) על G' שמתחיל בצומת s (שקיימת גם ב־G וגם ב־G'). שימו לב עתה שסדרת המ"מ עם האינדכסים הזוגיים $Y_0, Y_2, \ldots$ מתפלגת בדיוק כמו הילוך מקרי מהוסס על הגרף המקורי G, בעוד שהמ"מ במקומות האי־זוגיים $Y_1, Y_3, \ldots$ לעולם לא יקבלו ערכים המתאימים לצמתי הגרף המקורי, ובפרט לא יוכלו לקבל את t. מכאן שזמן הביקור הממוצע בין s ו־t לפי ההילוך המהוסס על G זהה בדיוק למחצית זמן הביקור הממוצע בין s ל־t לפי ההילוך הלא־מהוסס על G'. נסמן ב־$\hat{k}_{st}$ את זמן הביקור לפי ההילוך המהוסס, ב־k'_{st} את זמן הביקור לפי ההילוך על G', וב־k_{st} את זמן הביקור לפי הילוך לא מהוסס על G. כל שנותר עתה הוא לכתוב:

$$\hat{k}_{st} = \frac{1}{2}k'_{st} = m'R'_{st} = 4mR_{st} = 2k_{st}$$

לא חוזרים לאחור

השיטה הכי טובה היא לנתח את התפלגות של X_{i-1}, X_i ביחד כהתפלגות על זוגות סדורים של צמתים. נראה באינדוקציה על i שההתפלגות הזו (כשאינה מותנה על משתנים אחרים) היא יוניפורמית מעל קבוצת $2|E|$ האפשרויות עבור קשת מכוונת של G (לכל קשת uv של G יש שני כיוונים אפשריים, מ־u ל־v או מ־v ל־u). מזה נובע בפרט שההתפלגות הלא־מותנה של X_i היא זו המתוארת בשאלה.

הבסיס הוא הזוג X_0X_1, וזה נובע מההגדרה: לכל u, v שהם שני צמתי קצה של קשת של G, חישוב ישיר נותן $\Pr[X_0 = u \wedge X_1 = v] = \frac{d(v)}{2|E|} \cdot \frac{1}{d(v)} = \frac{1}{2|E|}$.

עבור המעבר, נניח שההנחה מתקיימת עבור $X_{i-1}X_i$, ונראה שהיא מתקיימת עבור X_iX_{i+1}. נניח ש־u, v הם שני צמתי קצה של קשת של G. אלו יכולים להיות הערכים של X_i, X_{i+1} אך ורק אם מתקיים $X_i = u$ אולם לא מתקיים $X_{i-1} = v$ (בגלל תנאי חוסר החזרה לאחור). ישנם $d(u) - 1$ שכנים של u ששונים מ־v. לכל שכן w כזה, לפי הנחת האינדוקציה $\Pr[X_{i-1} = w \wedge X_i = v] = \frac{1}{2|E|}$, ואלו מאורעות זרים. כמו כן, לכל w כזה מתקיים $\Pr[X_{i+1} = v | X_{i-1} = w \wedge X_i = u] = \frac{1}{d(u)-1}$, לפי הגדרת ההילוך חסר החזרות. מסיימים לפי נוסחת ההסתברות השלמה עבור ההתפלגות המותנה על המאורע $X_i = u$ (לא סוכמים על ערכים אפשריים של X_i):

$$
\begin{aligned}
\Pr[X_i = u \wedge X_{i+1} = v] &= \sum_{w \in N(V)} \Pr[X_{i+1} = v | X_{i-1} = w \wedge X_i = u] \Pr[X_{i-1} = w \wedge X_i = u] \\
&= (d(u) - 1) \cdot \frac{1}{d(u) - 1} \cdot \frac{1}{2|E|} = \frac{1}{2|E|}
\end{aligned}
$$

קשיים בהתקדמות

כיוון שכל צעד בהילוך אינו תלוי בצעדים הקודמים, t_k הוא גם תוחלת ההפרש בין ה־i שהוא הפעם ה־r עבורו $X_i = k - 1$ וה־$j > i$ הקטן ביותר עבורו $X_j = k$.

נחשב את t_k על סמך $t_1, \ldots, t_{k-1}$. נסמן ב־i את האינדקס הקטן ביותר עבורו $X_i = k - 1$, וננתח את ההתפלגות של j, האינדקס הקטן ביותר עבורו $X_j = k$, כאשר מתנים אותה על ערכו של i.

נסמן ב־i_r את האינדקס ה־r עבורו $x_{i_r} = k - 1$ (בפרט $i_1 = i$). בהסתברות $\frac{1}{3}$ בדיוק מתקיים $X_{i_1+1} = k$ ואז $j = i_1 + 1$, ובהסתברות $\frac{2}{3}$ מתקיים $X_{i_1+1} = k - 2$ ואז תוחלת i_2 היא $i_1 + 1 + t_{k-1}$. אז שוב בהסתברות $\frac{1}{3}$ נקבל $j = i_2 + 1$ ובהסתברות $\frac{2}{3}$ נקבל $X_{i_2+1} = k - 2$. באינדוקציה נקבל שההסתברות עבור $i_r < j < i_{r+1}$ היא $\frac{1}{3}(\frac{2}{3})^{r-1}$, וכאשר מתנים על המאורע הזה, תוחלת $i_r - i_1$ תהיה $(r - 1)(t_{k-1} + 1)$ ויתקיים $j = i_r + 1$. עתה משתמשים בנוסחת ההסתברות השלמה לקבלת נוסחת הנסיגה:

$$
\begin{aligned}
t_k &= \sum_{r=1}^{\infty} \mathrm{E}\left[j - i_1 | i_r < j < i_{r+1}\right] \cdot \Pr[i_r < j < i_{r+1}] \\
&= \sum_{r=1}^{\infty} (1 + (r - 1)(t_{k-1} + 1)) \cdot \frac{1}{3} (\frac{2}{3})^{r-1} \\
&= \frac{-t_{k-1}}{3} \sum_{r=1}^{\infty} (\frac{2}{3})^{r-1} + \frac{t_{k-1} + 1}{3} \sum_{r=1}^{\infty} (\frac{2}{3})^{r-1} r \\
&= 2t_{k-1} + 3
\end{aligned}
$$

פתרון נוסחת הנסיגה נותן $t_k = 3 \cdot 2^k - 3$.

טיול בגרף נאה

בגרף 3־רגולרי בעל n צמתים יש בדיוק $m = \frac{3}{2}n$ קשתות. עתה נחסום את ההתנגדות השקולה R_{st}. מכיוון שהגרף הוא 2־קשיר, קיימים בין s ל־t שני מסלולים זרים בצמתים. תוספת צמתים וקשתות יכולה רק להקטין את ההתנגדות השקולה, ולכן אפשר לחסום את ההתנגדות השקולה ע"י זו של שני המסלולים האלו בלבד. אם אורכיהם הם αn ו־βn בהתאמה, אז מהנוסחא עבור נגדים במקביל נקבל התנגדות שקולה של $\frac{\alpha\beta}{\alpha+\beta}n$. קצת אלגברה חוסמת את זה ע"י $\frac{\alpha+\beta}{4}n$, ומכיוון ש־$\alpha + \beta \leq 1$ (המסלולים הם זרי צמתים פרט ל־s ו־t) קיבלנו $R_{st} \leq \frac{1}{4}n$, ולכן $k_{st} = 2mR_{st} \leq \frac{3}{4}n^2$ כנדרש.

בלי הרבה נפנוף ידיים (חלק ראשון)

נראה שעבור כל ϵ קיים S כך שאם $s > S$ אז $\mathrm{E}[H_s] = (1 \pm \epsilon)s\pi_v$. ראשית נראה זאת עבור המקרה שבו הגרף אינו דו צדדי. נסמן ב־A_t את משתנה האינדיקטור שמקבל 1 אם $X_t = v$ ומקבל 0 אחרת. מכיוון שההתפלגות הלא מותנה של X_t שואפת להתפלגות הסטציונרית עבור $t \to \infty$, קיים T כך שאם $t > T$ אז $\mathrm{E}[A_t] = (1 \pm \epsilon/2)\pi_v$. כמו כן $0 \leq \mathrm{E}[A_t] \leq 1$ גם עבור $t \leq T$. לכן ניתן לבחור $S = 2T/\epsilon$ ואז עבור $s > S$ נקבל $\mathrm{E}[H_s] = \sum_{t=1}^{s} \mathrm{E}[A_t] = (1 \pm \epsilon)s\pi_v$ כנדרש.

עבור המקרה שבו הגרף דו צדדי, במקום לנתח את A_t מנתחים את $(A_t + A_{t+1})/2$. סכום זה ישאף להתפלגות הסטציונרית עבור $t \to \infty$, מכיוון שסכום ההתפלגויות הלא־מותנות של X_{t+1} ושל X_t לא יכיל בפירוקו לפי הווקטורים העצמיים של מטריצת המעבר את זה המתאים לערך העצמי -1 (המקדמים המתאימים יתקזזו).

בלי הרבה נפנוף ידיים (חלק שני)

נראה שעבור כל ϵ, δ קיים S כך שאם $s > S$ אז $H_s = (1 \pm \epsilon)s/\tau$ בהסתברות לפחות $1 - \delta$. נגדיר סדרה של מ"מ $Y_1, Y_2, \ldots$ אשר יקבעו ע"י ההילוך המקרי שלנו. Y_j יהיה מספר הצעדים בין הביקור ה־$j-1$ לבין הביקור ה־j ב־v, כאשר הביקור $X_0 = v$ יקרא "הביקור ה־0". נשים לב שלכל $j > 0$ מתקיים $\mathrm{E}[Y_j] = \tau$, שה־Y_j כולם ב"ת (עקב תכונת חוסר הזיכרון של ההילוך המקרי), ושקיים α סופי (תלוי בגרף) כך ש־$\mathrm{V}[Y_j] = \alpha$ (לא קשה להוכיח זאת, אבל בשאלה עצמה נאמר שלא צריך).

אם לא מתקיים $H_s = (1 \pm \epsilon)s/\tau$, אז ישנן שתי אפשרויות. אם $H_s > (1 + \epsilon)s/\tau$, אז בפרט חייב להתקיים $\sum_{j=1}^{(1+\epsilon)s/\tau} Y_j \leq s$. עתה אפשר להשתמש בשיטת המומנט השני: מתקיים $\mathrm{E}[\sum_{j=1}^{(1+\epsilon)s/\tau} Y_j] = (1 + \epsilon)s$ וכן $\mathrm{V}[\sum_{j=1}^{(1+\epsilon)s/\tau} Y_j] = (1 + \epsilon)s\alpha/\tau$, ולכן עבור s גדול דיו הסיכוי לסטיה של $\sum_{j=1}^{(1+\epsilon)/\tau} Y_j$ ב־ϵs מהממוצע חסום ע"י $\delta/2$.

אם $H_s < (1 - \epsilon)s/\tau$, אז בפרט חייב להתקיים $\sum_{j=1}^{(1-\epsilon)s/\tau} Y_j \geq s$, וגם כאן אפשר להשתמש בשיטת המומנט השני ולקבל שגם כאן עבור s גדול דיו הסיכוי לסטיה של $\sum_{j=1}^{(1-\epsilon)s/\tau} Y_j$ ב־ϵs מהממוצע חסום ע"י $\delta/2$. מאיחוד שתי האפשרויות לסטיה אנו מקבלים את המבוקש.

גם זו הרמונית

נסמן ב־$X_0, X_1, \ldots$ הילוך מקרי המתחיל ב־s, ועבור $v \in V$ ו־$k \geq 0$ נגדיר את מ"מ האינדיקטור $A_k^{(v)}$ אשר מקבל 1 אם $X_k = v$ ולא קיים $l \leq k$ עבורו $X_l = t$, ואחרת מקבל 0. שימו לב שבפרט $A_0^{(v)} = 1$ אם ורק אם $v = s$, וש־$A_k^{(t)} = 0$ לכל k. כמו כן $\sum_{k=0}^{\infty} A_k^{(v)}$ הוא בדיוק מספר הפעמים שההילוך ביקר ב־v לפני שהגיע לראשונה ל־t, ולכן מלינאריות התוחלת מתקיים עבור הפונקציה שלנו $\phi_{st}(v) = \frac{1}{d(v)} \sum_{k=0}^{\infty} \mathrm{E}[A_k^{(v)}]$.

שימו לב עתה שלכל $k > 0$, לכל $v \in V \setminus \{s,t\}$ ולכל סדרת ערכים $\langle b_u \rangle_{u \in V}$ (שמתקבלים בהסתברות חיובית) משתני האינדיקטור שלנו מקיימים $\Pr[A_k^{(v)} = 1 | \forall_{u \in V}\, A_{k-1}^{(u)} = b_u] = \sum_{u \in N(v)} \frac{1}{d(u)} b_u$ ישירות מהגדרות ההילוך המקרי, ולכן מתקיים $\mathrm{E}[A_k^{(v)}] = \sum_{u \in N(v)} \frac{1}{d(u)} \mathrm{E}[A_{k-1}^{(u)}]$. זה נכון גם עבור מקרי הקצה ש־s ו/או t נמצאים בשכנים של v. דבר נוסף לשים לב הוא שעבור $v \in V \setminus \{s,t\}$ מתקיים $\phi_{st}(v) = \frac{1}{d(v)} \sum_{k=1}^{\infty} A_k^{(v)}$ (כי $A_0^{(v)} = 0$), ואז ניתן לסיים:

$$\phi_{st}(v) = \frac{1}{d(v)} \sum_{k=1}^{\infty} \mathrm{E}[A_k^{(v)}] = \frac{1}{d(v)} \sum_{u \in N(v)} \frac{1}{d(u)} \left(\sum_{k=0}^{\infty} \mathrm{E}[A_k^{(u)}] \right) = \frac{1}{d(v)} \sum_{u \in N(v)} \phi_{st}(u)$$