תרגיל ראשון

משחקים במספרים

עבור הסעיף הראשון, פשוט נחשב לכל בחירה של שני מספרים $k < l$ את סיכויי הזכיה של השחקן השני: בהסתברות $\frac{1}{2}$ הוא יראה את k, ובהינתן שזה קרה הוא יענה את התשובה הנכונה ("זה המספר הקטן") בהסתברות $\frac{1}{k}$. בדומה, בהסתברות $\frac{1}{2}$ הוא יראה את l ובהינתן מאורע זה הוא יענה נכון בהסתברות $1 - \frac{1}{l}$. סה"כ ההסתברות לתשובה נכונה שלו הוא $\frac{1}{2}\frac{1}{k} + \frac{1}{2}\frac{l-1}{l} = \frac{l+kl-k}{2kl} > \frac{1}{2}$.

עבור הסעיף השני, שימו לב שתיאור מלא של האסטרטגיה של השחקן השני תמיד מתמצא בזה שלכל מספר טבעי k השחקן יכריז עליו במידה ויראה אותו בהסברות p_k כמספר הגדול מהשניים, כאשר $p_1, p_2, \ldots$ תלויים באסטרטגיה הספציפית. מכיוון שכל ההסתברויות הנ"ל הן בין 0 ל־1, קיים j שעבורו $p_{j+1} - p_j < 2\epsilon$. לכן אם השחקן הראשון בוחר את זוג המספרים $j, j+1$, אז סיכוי הניצחון של השחקן השני נתון ע"י $\frac{1}{2}(1 - p_j) + \frac{1}{2}p_{j+1} = \frac{1}{2} + \frac{1}{2}(p_{j+1} - p_j) < \frac{1}{2} + \epsilon$.

פרדוקס יום ההולדת

הניתוח עבור $l = o(\sqrt{k})$: לכל $1 \leq i < j \leq l$ יהי $E_{i,j}$ המאורע ש־$X_i = X_j$. לכן, הסיכוי שיש שני מספרים זהים בתוך $X_1, \ldots, X_l$ חסום ע"י $\sum_{1 \leq i < j \leq l} \Pr[E_{i,j}] = \binom{l}{2} \cdot \frac{1}{k} = O(l^2/k)$. כאשר $l = o(\sqrt{k})$ אז החסם הנ"ל הוא $o(1)$ כנדרש.

הניתוח עבור $l = \omega(\sqrt{k})$: נניח ש־l זוגי (אחרת נשתמש ב־$l-1$). אנו נחסום למעשה את הסיכוי ש־$X_{l/2+1}, \ldots, X_l$ מכילים ערך אחד (או יותר) שהופיע ב־$X_1, \ldots, X_{l/2}$, מותנה על הסיכוי ש־$X_1, \ldots, X_{l/2}$ קיבלו ערכים שונים זה מזה. מספיק לחסום את ההסתברות המותנה הנ"ל מכיוון שאם $X_1, \ldots, X_{l/2}$ לא קיבלו ערכים שונים זה מזה אז ממילא יש לנו כבר ערך שהתקבל פעמיים.

מכיוון שכל X_i עבור $i > l/2$ נבחר באופן ב"ת לחלוטין בערכי $X_1, \ldots, X_{l/2}$ (שהרי כל הערכים נבחרו באופן ב"ת), אז ההסתברות המותנה ש־X_i קיבל את אחד מ־$l/2$ הערכים הנ"ל היא $\frac{l}{2k}$. לכן, ההסתברות שאף X_i עבור $l/2 < i \leq l$ לא קיבל ערך מהנ"ל היא $(1 - \frac{l}{2k})^{l/2} < e^{-l^2/4k}$. כאשר $l = \omega(\sqrt{k})$ אז החסם הנ"ל הוא $e^{-\omega(1)} = o(1)$.

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 $f(u),f(v)$.

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.

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 $p=1$.

תרגיל שני

למת הבידוד עם זוגיות

ראשית נשים לב שבסיכוי $1-O(\frac{n}{m})$ קורה הדבר הבא: גם קיימת קבוצה יחידה $F\in\mathcal{F}$ עם משקל מינימלי, וגם אף איבר ב־A לא מקבל את המשקל 1. בפרט מספר ההשמות המקיימות זאת הוא $(1-O(\frac{n}{m}))n^m$. נגדיר עתה את הטרנספורמציה הבאה מקבוצת כל ההשמות האפשריות של w המקיימות את התנאים הנ"ל, לתוך קבוצת כל ההשמות שעבורן יש קבוצה יחידה עם משקל מינימלי אי־זוגי, באופן הבא:

תהיא F הקבוצה היחידה עם משקל מינימלי. יהי a האיבר עם האינדקס המינימלי ב־F. לקבלת w', אם $w(F)$ כבר אינו זוגי, אז לא נעשה כלום, אבל אם $w(F)$ זוגי אז נוריד את $w(a)$ ב־1. שימו לב שעדיין F היא הקבוצה היחידה עם משקל מינימלי, וכן שאחרי הטרנספורמציה w' תהיה השמה חוקית (כי קודם לכן לא היו איברים עם משקל 1) ושעבורה המשקל המינימלי הוא אי־זוגי.

עתה נשים לב שהעתקה זו היא לכל היותר שניים לאחד: השמה w' יכולה להתקבל או מ־w ששווה לה, או מ־w המתקבלת ע"י תוספת 1 לאיבר עם האינדקס המינימלי בקבוצה עם המשקל המינימלי (יש גם מקרים שבהם w' יכול להתקבל רק מעצמו, אולם מקרים כאלו רק מגדילים את החסם על הגודל של קבוצת ה־w' האפשריות). לכן בהכרח מספר ההשמות האפשריות w' הוא לפחות $\frac{1}{2}(1-O(\frac{n}{m}))n^m$, ז"א שהשמה מקרית לפונקצית המשקל תתן השמה כנ"ל בסיכוי לפחות $\frac{1}{2}-O(\frac{n}{m})$.

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 שצמתים 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 $\mathrm{V}[X]$.

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 $\Pr\left[|X-\frac{1}{4}\alpha n^2|>n^{7/4}\right]<n^{-1/2}$, ז"א שבסיכוי $1-o(1)$ מתקיים $X=\frac{1}{4}\alpha n^2\pm n^{7/4}=\frac{1}{4}\alpha n^2+o(n^2)$ כנדרש.

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 קרו 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 ניתוח, 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 (אלא אם כן הוא כבר עצר 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 התעלמתי כאן מקבועים מדויקים). נסמן לבסוף ב־T את המ"מ שמקבל את מספר הצעדים שלקח לאלגוריתם לעצור, ונקבל:

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

תרגיל שלישי

קיום תת גרף ספציפי בגרף צפוף

ניתן להוכיח את המבוקש עם $c = 10$, 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 $\Pr[\neg E_{uv}] < \frac{1}{10d}$, 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. להמשך הטיעון אבל 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 ש־G חייב להיות עותק של H כתת־גרף.

תתי גראפים בגראפים צפופים ־ הדור הבא

נניח שההגרלה של תת הקבוצה של הצמתים נעשית באמצעות הגרלה של פונקציה חד־חד ערכית $f : \{1, \ldots, \frac{n}{2}\} \to V(G)$. נסמן ב־$X$ את המשתנה המקרי של מספר הקשתות המושרה על קבוצת ערכי f, ונגדיר עתה את המרטינגל $X_0, \ldots, X_{n/2}$ כמרטינגל החשיפה של X לפי תתי התחומים $\mathcal{D}_i = \{1, \ldots, i\}$. מכיוון שערכי f אינם מוגרלים באופן בלתי תלוי, אי אפשר לחסום את $|X_i - X_{i-1}|$ באמצעות הקריטריון הקומבינטורי, אלא יש להוכיח את החסם ישירות (עם זאת, העובדה שזהו מרטינגל אינה תלויה בדרך שהגרלנו את f).

אנו נראה את הדבר הבא: לכל i, לכל סדרת צמתים ללא חזרות $u_1, \ldots, u_i$ ולכל u שאינו נמצא בסדרה זו מתקיים

$$\left|\mathrm{E}[X|f(1) = u_1, \ldots, f(i) = u_i] - \mathrm{E}[X|f(1) = u_1, \ldots, f(i-1) = u_{i-1}, f(i) = u]\right| \leq \frac{1}{2}n$$

מכאן ינבע שמתקיים (ע"י כך שמציבים את $u_1, \ldots, u_i$ להיות i הערכים הראשונים של $\tilde{f}$ שהוגרלה ע"פ ההגדרה של מרטינגל חשיפה) ש־$|X_i - X_{i-1}|$ חסום ע"י

$$\frac{1}{n+1-i} \sum_{u \notin \{u_1, \ldots, u_{i-1}\}} \left|\mathrm{E}[X|f(1) = u_1, \ldots, f(i) = u_i] - \mathrm{E}[X|f(1) = u_1, \ldots, f(i-1) = u_{i-1}, f(i) = u]\right|$$

שחסום ע"י $\frac{1}{2}n$, ולאחר זאת הפעלה ישירה של משפט אזומה נותנת את המבוקש.

לסיום ההוכחה אם כן ניקח $u_1,\ldots,u_i,u$ כמקודם, ונגדיר התאמה חד־חד ערכית ועל מקבוצת הפונקציות המקיימות $f(1)=u_1,\ldots,f(i)=u_i$ אל קבוצת הפונקציות המקיימות $f(1)=u_1,\ldots,f(i-1)=u_{i-1},f(i)=u$. אם f היא פונקציה מהקבוצה הראשונה, אז ישנם שני מקרים. במקרה הראשון קיים j כך ש־$f(j)=u$, ולפי הנתונים חייב אז להתקיים $j>i$. במקרה כזה, לקבלת f' נחליף את הערכים $f'(j)=f(i)$ ו־$f'(i)=f(j)$ (ושאר ערכי f' יהיו זהים לאלו של f). במקרה השני לא קיים j עבורו $f(j)=u$, ואז פשוט נקבע את $f'(i)$ להיות u בעוד שאר הערכים ישארו ללא שינוי.

עתה נשווה את הערך של X לפי f ולפי f'. במקרה הראשון הוא ללא שינוי כלל (קבוצת הצמתים נשארה זהה, רק הסדר השתנה), ובמקרה השני הוא השתנה בלא יותר מ־$\frac{1}{2}n$, שזה ההפרש בין מספר השכנים האפשרי בין הצומת $f(i)$ או $f'(i)$ לבין קבוצת שאר הצמתים $\{f(j):j\neq i\}$. מכיוון שכל פונקציה בקבוצת היחוס המתאימה נבחרת בהסתברות שווה (ההגרלה המקורית הינה יוניפורמית מעל קבוצת הפונקציות המותרות), נובע מקיום ההתאמה הנ"ל שאכן מתקיים

$$\left|\mathrm{E}[X|f(1)=u_1,\ldots,f(i)=u_i]-\mathrm{E}[X|f(1)=u_1,\ldots,f(i-1)=u_{i-1},f(i)=u]\right|\leq\frac{1}{2}n$$

קבוצה קשירה על הקוביה

אנו נראה לכל $\epsilon>0$ שעבור n גדול דיו ועבור כל קבוצה A ניתן לשנות אותה בלא יותר מ־$\epsilon 2^n$ מקומות לקבלת המבוקש. את זאת נעשה בשלושה שלבים. בשלב הראשון, נוסיף ל־A את כל הווקטורים ב־$V=\{0,1\}^n$ שעבורם מספר האחדות קטן מ־$\frac{1}{3}n$. עבור n גדול דיו לא קשה לראות שמספר האיברים הנ"ל קטן מ־$\frac{1}{3}\epsilon 2^n$. השלב השני הוא מקרי: לכל איבר עם לפחות $\frac{1}{3}n$ אחדות, נוסיף אותו ל־A (אם לא היה שם ממילא) בהסתברות $\frac{1}{3}\epsilon$ באופן בלתי תלוי באיברים האחרים. נסמן את הקבוצה לאחר שתי ההוספות ב־A'.

עתה נגיד שווקטור $\underline{x}$ (אם הוא עצמו ב־A' או לא) בעל $j\geq\frac{n}{3}$ אחדות בדיוק הוא בעל מסלול יורד אם קיימים $\{\underline{x}^{(i)}:\frac{n}{3}\leq i<j\}$ שכולם ב־A', כך שלכל i מספר האחדות ב־$\underline{x}^{(i)}$ הוא i בדיוק וכן $\underline{x}^{(i)}$ ו־$\underline{x}^{(i-1)}$ נבדלים ביניהם בקורדינטה אחת בדיוק, ובנוסף $\underline{x}^{(j-1)}$ ו־$\underline{x}$ נבדלים בקורדינטה אחת בדיוק. נחסום את הסיכוי של־$\underline{x}$ אין מסלול יורד: הסיכוי שלא קיים $\underline{x}^{(j-1)}$ אשר מקיים את הדרישות הוא בבירור לכל היותר $(1-\frac{1}{3}\epsilon)^{n/3}$. בהינתן שקיים $\underline{x}^{(j-1)}$ כנדרש נבחר אחד כזה שרירותית, ונחשב את ההסתברות המותנה שלא קיים $\underline{x}^{(j-1)}$ מתחתיו הנבדל ממנו בקורדינטה. מכיוון שההגרלות בשכבת האיברים עם $j-2$ אחדות הן ב"ת בהגרלות בשכבה הקודמת, שוב יש לנו חסם של $(1-\frac{1}{3}\epsilon)^{n/3}$, גם כאשר מתנים את זה במאורע הקודם. משיקולים דומים אפשר להמשיך לחסום את ההסתברות גם לשכבות הבאות, ולקבל שעבור n גדול דיו הסיכוי ש־$\underline{x}$ הוא חסר מסלול יורד חסום ע"י $\frac{1}{3}\epsilon$.

בשלב השלישי פשוט נסיר מ־A' את כל האיברים שעבורם אין מסלול יורד. שימו לב שההגדרה של מסלול יורד היא כזו שבהסרה כנ"ל לא נפגע במסלולים יורדים של איברים מ־A' שלא הוסרו. לכן, הקבוצה המתקבלת לאחר ביצוע שלב זה היא קשירה (זיכרו שהוספנו בשלב הראשון את כל האיברים עם פחות מ־$\frac{1}{3}n$ אחדות). כמו כן, תוחלת מספר האיברים שהוסרו כאן חסומה ע"י $\frac{1}{3}\epsilon 2^n$.

לסיכום, תוחלת מספר האיברים שהוספו או הורדו בשני השלבים האחרונים חסומה ע"י $\frac{2}{3}\epsilon 2^n$, ולכן יש אפשרות לבצע אותם בצורה שאכן פחות ממספר זה של איברים השתנו. בתוספת החסם על מספר האיברים שהוספו בשלב הראשון, קיבלנו את המבוקש.