הרצאה 7

1 מבוא לאלגוריתמים מקריים

אלגוריתם מקרי לעיתים יהיה פשוט או יעיל יותר מאלגוריתם דטרמיניסטי. נתאר שתי דוגמאות:

1.1 שוויון פולינומים

בהינתן שני פולינומים: אחת בצורת מכפלה והשנייה בצורה קנונית:

$$F(x) = \prod_{i=1}^{d}(x - a_i)\,, \qquad G(x) = x^d + \sum_{j=0}^{d-1} b_j x^j$$

נרצה לבדוק האם $F(x) \equiv G(x)$. דרך ישירה לעשות את זה היא להעביר את $F(x)$ לצור קנונית ולהשוות מקדמים.
אבל זה ידרוש $\Theta(d^2)$ פעולות כפל וחיבור. האלגוריתם המקרי הבא יהיה הרבה יותר מהיר:

אלגוריתם מקרי להשוואת פולינומים:

קלט: פולינומים $F(x), G(x)$ כמתואר לעיל.
פלט: $F(x) \equiv G(x)$ או $F(x) \not\equiv G(x)$.

1. נבחר שלם $r \in \{1,2, \dots 100d\}$ באופן מקרי ואחיד.

2. אם $F(r) \neq G(r)$, נחזיר $F(x) \not\equiv G(x)$. אחרת נחזיר $F(x) \equiv G(x)$.

סיבוכיות האלגוריתם: נניח שהבחירה בשלב 1 דורשת זמן קבוע $O(1)$. החישוב של $F(r)$, $G(r)$ דורש זמן $\Theta(d)$.
שזה סה"כ זמן הריצה של האלגוריתם.

ננתח את הנכונות: נניח ש $F(x) \equiv G(x)$. במקרה הזה, בוודאות $F(r) \neq G(r)$ ולכן נחזיר את התשובה הנכונה.
נניח עכשיו ש $F(x) \not\equiv G(x)$. אם $F(r) \neq G(r)$, נחזיר תשובה נכונה. אם $F(r) = G(r)$, נחזיר תשובה שגויה.
מה ההסתברות שזה יקרה?

אם $F(x) \not\equiv G(x)$, אז הפולינום $H(x) = F(x) - G(x)$ הוא לא פולינום האפס ולכן יש לכל היותר d שורשים.
ובפרט, לכל היותר d שורשים בקבוצה $\{1,2, \dots 100d\}$.
בגלל שבחרנו את r מתוך הקבוצה באופן אחיד, נובע ש $\mathbb{P}\big(F(r) = G(r)\big) \leq \frac{1}{100}$.

לסיכום, אם האלגוריתם החזיר שהם לא שווים, זה נכון.
אם האלגוריתם החזיר שהם שווים, זה נכון בהסתברות לפחות 0.99, מהיר יותר אבל אולי שגוי.
נרצה להקטין מאד את הסיכוי לטעות, לדוגמה על ידי הגדלת הקבוצה שמתוכה נבחר את r.
אבל זה עלול להשפיע על הסיבוכיות (הסיבוכיות של בחירה מתוך קבוצה יכולה להיות תלויה בגודל הקבוצה).
בנוסף, עבודה עם מספרים גדולים מאד יכולה להיות בעייתית עבור מחשבים.
ניתן פתרון יותר כללי ויותר טוב:

אלגוריתם מקרי משופר להשוואת פולינומים:

קלט: פולינומים $F(x), G(x)$ כמתואר לעיל, ומספר שלם חיובי k.
פלט: $F(x) \equiv G(x)$ או $F(x) \not\equiv G(x)$.

1. עבור $1 \leq i \leq k$:

a. נבחר שלם $r_i \in \{1,2, \dots 2d\}$ באופן מקרי ואחיד.

b. אם $F(r_i) \neq G(r_i)$, נעצור נחזיר $F(x) \not\equiv G(x)$.

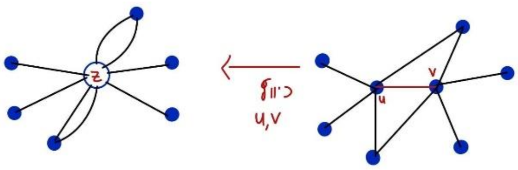
2. נחזיר $F(x) \equiv G(x)$.

בניתוח דומה למה שעשינו מקודם, נקבל שהסיבוכיות היא $\Theta(kd)$ (שזה $\Theta(d)$ אם k הוא קבוע). בנוסף, אם האלגוריתם החזיר שהם שונים, זה בוודאות נכון. ואם האלגוריתם החזיר שהם שווים, זה נכון בהסתברות לפחות $1 - 2^{-k}$.

1.2 אלגוריתם חתך מינימלי רנדומלי

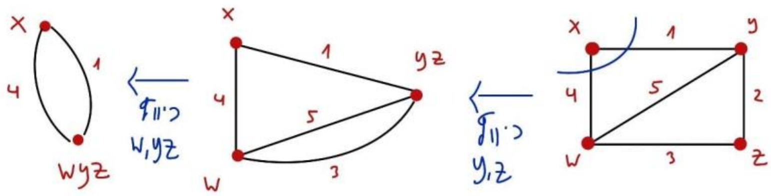
יהי $G = (V, E)$ גרף קשיר. **חתך** בG הוא קבוצה $A \subseteq E$ כך ש $G \backslash A$ לא קשיר.
נרצה למצוא חתך בגודל מינימלי. נתאר אלגוריתם מקרי פשוט, שמשתמש ב**כיווץ קשתות:**

הגדרות: **מולטיגרף** הוא גרף שמאפשר קיום של יותר מצלע אחת בין שני קודקודים (צלעות מקבילות).
לולאה היא צלע מקודקוד לעצמו.



בהינתן מולטיגרף בלי לולאות, **כיווץ** צלע uv נעשה על ידי איחוד u,v לקודקוד חדש z_{uv} ומחיקת כל הצלעות בין u ל-v. כל צלע שחיברה בין u או v לקודקוד אחר, עכשיו תחבר בין אותו קודקוד לקודקוד החדש z_{uv}. נקרא לגרף החדש $G \backslash uv$ ונשים לב שיכולות להיות בו צלעות מקבילות אבל אין בו לולאות.

אלגוריתם מקרי למציאת חתך מינימלי:



קלט: גרף קשיר G על n קודקודים.
פלט: חתך של G.

1. יהי $G_0 = G$ מולטיגרף.
2. עבור $1 \leq i \leq n-2$:
   a. נבחר צלע $e_i \in E(G_{i-1})$ באופן מקרי ואחיד.
   b. נגדיר את $G_i = G_{i-1} \backslash e_i$ (נכווץ את הצלע).
3. נחזיר את $E(G_{n-2})$.

מכיוון שכל צעד בלולאה לוקח זמן $O(n)$, זמן הריצה הכולל הוא $O(n^2)$.

הפלט תמיד יהיה 2 קודקודים עם צלעות ביניהן, והצלעות האלה מהוות חתך.

טענה 1.1: האלגוריתם מחזיר חתך מינימלי בהסתברות לפחות $\frac{2}{n(n-1)} = \binom{n}{2}^{-1}$. (זו לא הסתברות גבוהה).

הוכחה: יהי $A \subseteq E(G)$ חתך מינימלי כלשהו. נגדיר את k הגודל של החתך.
לכל $1 \leq i \leq n-2$, נגדיר את E_i המאורע $e_i \notin A$. כלומר בסיבוב הזה, לא כיווצנו צלע ששייכת לחתך.
נשים לב ש:

$$\mathbb{P}(\text{the algorithm returns a min}-\text{cut}) \geq \mathbb{P}(\text{the algorithm returns A}) \overset{\aleph}{=} \mathbb{P}\left(\bigcap_{j=1}^{n-2} E_j\right)$$

$$= \mathbb{P}(E_1) \cdot \mathbb{P}(E_2 | E_1) \cdot \dots \cdot \mathbb{P}\left(E_{n-2} \middle| \cap_{i=1}^{n-3} E_i\right)$$

א – כל פעם שמכווצים צלע, היא נמחקת. כלומר החתך זה כל הצלעות שלא מחקנו. אז אם החזרנו את A זה אומר שבכל שלב בחרנו צלע שלא שייכת ל-A.

ואכן, מכיוון ש A הוא חתך, אם נוריד אותו זה מחלק את V לשתי קבוצות: S, $V \backslash S$. שאין צלעות ביניהן.
אם נכווץ צלעות ששני הקודקודים שלהם שייכים לאחת הקבוצות $n-2$ פעמים,
אז S ו- $V \backslash S$ יהפכו כל אחת לקודקוד והצלעות שמחברות בין שני הקודקודים האלה הן בדיוק A.
נשאר להוכיח ש $\mathbb{P}\left(\cap_{j=1}^{n-2} E_j\right) \geq \binom{n}{2}^{-1}$.
מכיוון שהחתך המינימלי הוא בגודל k, כל חתך הוא בגודל לפחות k.
בפרט, הדרגה המינימלית היא לפחות k (כי אחרת נוכל פשוט לקחת רק את הקודקוד הזה).
ולכן, $|E(G)| \geq kn \backslash 2$. (הדרגה המינימלית כפול מספר הקודקודים, חלקי 2)
עכשיו נוכל לקבל את ההסתברות של E_1: מכיוון ש e_1 נבחרה באופן מקרי ואחיד מתוך $E(G)$, נקבל ש:

$$\mathbb{P}(E_1) = \frac{|E(G) \backslash A|}{|E(G)|} = \frac{|E(G)| - |A|}{|E(G)|} = 1 - \frac{|A|}{|E(G)|} \geq 1 - \frac{k}{kn \backslash 2} = 1 - \frac{2}{n}$$

נשים לב ש $|V(G_1)| = n-1$, ושהגודל של חתך מינימלי ב-G_1 הוא לפחות k
(כי חתך מינימלי של G_1 הוא גם חתך מינימלי של G). ולכן:

$$\mathbb{P}(E_2|E_1) \geq \frac{|E(G_1)\backslash A|}{|E(G_1)|} = \frac{|E(G_1)| - |A|}{|E(G_1)|} = 1 - \frac{|A|}{|E(G_1)|} \geq 1 - \frac{k}{k(n-1)\backslash 2} = 1 - \frac{2}{n-1}$$

ובאופן דומה, לכל $3 \leq i \leq n-2$:

$$\mathbb{P}\left(E_i \middle| \cap_{j=1}^{i-1} E_j\right) \geq 1 - \frac{k}{k(n-i+1)\backslash 2} = 1 - \frac{2}{n-i+1}$$

(1 פחות k חלקי החסם התחתון על מספר הצלעות).

בסה"כ נקבל ש:

$$\mathbb{P}\left(\bigcap_{j=1}^{n-2} E_j\right) = \mathbb{P}(E_1) \cdot \mathbb{P}(E_2|E_1) \cdot \dots \cdot \mathbb{P}\left(E_{n-2} \middle| \cap_{i=1}^{n-3} E_i\right) \geq \prod_{i=1}^{n-2}\left(1 - \frac{2}{n-i+1}\right) = \prod_{i=1}^{n-2}\left(\frac{n-i+1-2}{n-i+1}\right)$$
$$= \prod_{i=1}^{n-2}\left(\frac{n-i-1}{n-i+1}\right) = \left(\frac{n-2}{n}\right)\left(\frac{n-3}{n-1}\right)\left(\frac{n-4}{n-2}\right) \cdot \dots \cdot \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{1}{3}\right) =^{\text{א}} \frac{2}{n(n-1)} = \binom{n}{2}^{-1}$$

כמו באלגוריתם של הפולינומים, אם נחזור על האלגוריתם מספיק פעמים נקבל הסתברות גבוהה לתשובה נכונה.
לדוגמה, אם נריץ $n(n-1)\ln n$ פעמים ונחזיר את החתך הכי קטן שמצאנו, נקבל שההסתברות לטעות היא לכל היותר:

$$\left(1 - \frac{2}{n(n-1)}\right)^{n(n-1)\ln n} \leq^{\text{א}} e^{-2\ln n} = \frac{1}{n^2}$$

א - $1 - p \leq e^{-p}$

הסיבוכיות היא $O(n^4 \ln n)$