אלגוריתמים 2 – עבודה 3

שאלה 1

א. יהי MST עפ״מ של G.
נניח בשלילה שקיים T ע״פ של G כך ש $\lambda(T) < \lambda(MST)$.
תהי $e \in MST$ כך ש $w(e) = \lambda(MST)$.
נסתכל על החתך $(S, \bar{S})$ המתקבל ע״י $MST \backslash \{e\}$.
תהי $e' \in T$ כך ש $(u, v) = e'$ כך ש $u \in S$ וגם $v \in \bar{S}$.
ע״פ משפט שראינו בכיתה $MST \backslash \{e\} \cup \{e'\}$ הינו ע״פ של G.
נחלק למקרים:
I. $w(e) \le w(e')$: אזי $\lambda(MST) = w(e) \le w(e') \le \lambda(T)$ בסתירה להנחה.
II. $w(e) > w(e')$: אזי

$$\sum_{x \in MST} w(x) = \sum_{x \in MST \backslash \{e\}} w(x) + w(e) > \sum_{x \in MST \backslash \{e\}} w(x) + w(e') = \sum_{x \in MST \backslash \{e\} \cup \{e'\}} w(x)$$

בסתירה לכך ש MST הינו עפ״מ של G.

ב. **תיאור האלגוריתם**

1. $E_T = \emptyset, G' = (V, E'), E' = E, V = V'$
2. $while(E' \neq \emptyset)$:
   2.1 $e_m = Select\left(\frac{|E'|}{2}, E'\right)$
   2.2 $G_m = (V, E_m), E_m = \{e \in E' \mid w(e) < w(e_m)\}$
   2.3 $if(G_m \text{ is connected})$:
      2.3.1 $G' = G_m$
      2.3.2 $jump \ to \ line \ 2$
   2.4 $else$:
      2.4.1 $for(C \text{ in } components(G_m))$:
         2.4.1.1 $E_T = E_T \cup SpanningTree(C)$
      2.4.2 $G' = Compress(G_m)$
3. $return \ T = (V, E_T)$

הערות:

1. בשורה 2.1 נאתחל את E_m כך ש $|E_m| = \frac{|E'|}{2}$, במידה והאלגוריתם יצטרך לבחור בין מספר צלעות שוות משקל, הוא יגריל ביניהן.
2. בשורה 2.4.2 הפונקציה $Compress$ מכווצת כל רכיב קשירות ב G_m לכדי קודקוד יחיד, ומחברת כל שני קודקודים חדשים כאלה ב E' החדש בעזרת הצלע המינימלית המחברת בין קודקודיהם ב E' הישן.

הוכחת נכונות

טענה ראשית: האלגוריתם מחזיר $T = (V, E_T)$ ע״פ של G בעל $\lambda(T)$ מינימאלי.

טענה נשמרת: בסוף כל איטרציה של האלגוריתם, קיים $T' = (V, E_{T'})$ ע״פ כלשהו של G בעל $\lambda(T')$ מינימאלי, כך ש $E_T \subseteq E_{T'}$ ולכל $e \in E$ שהוסרה או כווצה עד לסוף איטרציה זו מתקיים $e \notin E_{T'}$.

הוכחה טענה ראשית:
בסוף ריצת האלגוריתם לכל $e \in E$ מתקיים כי $e \in E_T$ או ש e הוסרה או כווצה. לכן, ע״פ הטענה הנשמרת $T = (V, E_T)$ הינו ע״פ של G בעל $\lambda(T)$ מינימאלי.

הוכחת טענה נשמרת:
נוכיח באינדוקציה על מספר האיטרציות k.
בסיס $k = 1$:
I. G_m קשיר: אזי קיים $MST = (V, E_{MST})$ עפ״מ של G כך ש $E_{MST} \subseteq E_m$, לכן עבור כל $e \in E$ שהוסרה באיטרציה זו מתקיים $e \notin E_{MST}$.
ע״פ סעיף א׳ מתקיים כי MST הינו ע״פ של G בעל $\lambda(MST)$ מינימאלי, ובנוסף $E_T = \emptyset \subseteq E_{MST}$.
בסה״כ הראינו כי MST הינו ע״פ של G המקיים את תנאי הטענה.
II. G_m אינו קשיר: נסתכל על ריצת אלגוריתם קרוסקל עד לבחינת צלע e_m, משמע כאשר האלגוריתם עבר על מחצית הצלעות בעלות המשקל הנמוך.
נשים לב כי היער שנבנה ע״י קרוסקל מורכב מעפ״מים של רכיבי הקשירות של G_m.
נסמן את העץ שיחזור מריצה זו של קרוסקל כ MST.
נבחין כי קיימת ב MST צלע e' המקיימת $w(e) \leq w(e') = \lambda(MST)$ לכל $e \in E_m$.
נבנה $T' = (V, E_{T'})$ ע״פ של G כך שצלעותיו הן צלעות MST שאינן ב G_m והצלעות של הע״פים של רכיבי הקשירות שנבחרו ע״י האלגוריתם בשורה 2.4.1.1.
נבחין כי אכן מדובר בע״פ של G משום שהחלפנו תתי ע״פים של ה MST בתתי ע״פים אחרים.
בנוסף, מכיוון ש e' הינה צלע ב T' מתקיים כי $\lambda(MST) = \lambda(T')$.
ע״פ סעיף א׳, קיבלנו כי T' הינו ע״פ של G בעל $\lambda(T')$ מינימאלי, משמע T' הינו ע״פ של G המקיים את תנאי הטענה.
הנחת האינדוקציה:
נניח כי בסיום האיטרציה ה k של האלגוריתם קיים T' ע״פ של G המקיים את תנאי הטענה.
צעד האינדוקציה:
נסתכל על E_T בסוף האיטרציה ה $k+1$ ועל T' הע״פ של G המובטח לנו מהנחת האינדוקציה.
I. G_m קשיר: אזי ניתן לומר כי T' אינו מכיל צלעות שהוסרו באיטרציה זו שכן משקלן כבד ממשקל הצלעות שהוסרו (למעט המקרה של צלעות שוות משקל אותן ניתן להחליף ע״פ משפט החתך), ולא יתכן כי T' מכיל אחת מהן שכן הוא בעל $\lambda(T')$ מינימאלי. מכיוון שלא נוספו צלעות ל E_T במקרה זה ניתן לומר כי T' הינו ע״פ של G המקיים את תנאי הטענה (או ע״פ אחר בעל משקל זהה אליו ניתן להגיע ע״פ משפט החתך).
II. G_m אינו קשיר: באופן זהה לבסיס האינדוקציה ע״פ בחינה של ריצת קרוסקל.

זמן ריצה

בכל איטרציה מפעילים פונקציית $Select$ בזמן $O(|E'|)$, מגדירים את G_m בזמן $O(|E'|)$, בדיקת קשירות G_m בזמן $O(|E'|)$, ובחלק מהאיטרציות בניית עפ״מים של רכיבי G_m בזמן $O(|E'|)$ וצמצום G' בזמן $O(|E'|)$. בסה״כ ניתן לראות כי זמן ריצה של כל איטרציה הינו $O(|E'|)$.

סדרת $|E'|$ הינה $|E|, \frac{|E|}{2}, \frac{|E|}{4}, \ldots, 1$, ולכן זמן ריצת האלגוריתם הינו:

$$O\left(|E| + \frac{|E|}{2} + \frac{|E|}{4} + \cdots + 1\right) = \boldsymbol{O}(|\boldsymbol{E}|)$$

שאלה 2

<u>טענה :</u> אלג׳ $Boruvka$ עם כיווצים הינו אלג׳ דטרמיניסטי המוצא בזמן ליניארי MST בגרף מישורי.

<u>הוכחה :</u> ראשית, $Boruvka$ הינו אלג׳ דטרמיניסטי והוכחנו בכיתה כי הוא מחזיר MST.

כעת נראה כי עבור גרף מישורי G האלגוריתם רץ בזמן $O(n)$.

יהי גרף מישורי $G = (V,E)$, כאשר $|E| = n, |V| = m$.

ממישוריות G נקבל כי $m < 3n$, כלומר $m = O(n)$.

<u>אבחנה :</u> גרף מישורי אשר הוסרו ממנו צלעות או כווצו בו קודקודים נשאר מישורי.

לכן לכל j, באיטרציה ה-j של האלגוריתם הגרף $G_j = (V_j, E_j)$ שנקבל הוא מישורי (מאחר והפעולות היחידות שאנחנו עושים הן כיווץ קודקודים והסרת צלעות).

נסמן $n_j = |E_j|,\ m_j = |V_j|$.

כעת נרצה לחסום את זמן ריצת האלגוריתם $T(G)$:

$$\boldsymbol{T(G)} =^{(1)} \sum_{j=1}^{\log n} O(m_j) =^{(2)} \sum_{j=1}^{\log n} O(n_j) =^{(3)} \sum_{j=1}^{\log n} O\left(\frac{n}{2^j}\right) = O\left(\sum_{j=1}^{\log n} \frac{n}{2^j}\right) = \boldsymbol{O(n)}$$

(1) ראינו בכיתה כי זמן הריצה של כל איטרציית $Boruvka$ הינו $O(m_j)$, וכי מספר האיטרציות של האלגוריתם חסום ע״י $logn$.
(2) ע״פ האבחנה כי לכל j, הגרף $G_j = (V_j, E_j)$ שנקבל הינו מישורי.
(3) ראינו בכיתה כי לכל j, הגרף $G_j = (V_j, E_j)$ בעל לכל היותר $\frac{n}{2^j}$ קודקודים.

שאלה 3

תיאור האלגוריתם

1. $C = DFS(G) \ \backslash\backslash\ C\ is\ G's\ group\ of\ components$
2. $for\ c\ in\ C$:
    2.1 $for\ v\ in\ c$:
        2.1.1 $if(|N(v)| < |c| - 1): return\ true$
3. $return\ false$

הוכחת נכונות

טענה ראשית: האלג׳ מחזיר $true \Leftrightarrow G$ בעל תת גרף מושרה P_3.

טענת עזר: G בעל תת גרף מושרה $P_3 \Leftrightarrow$ קיים $c \in C$ כך שקיים $v \in c$ כך ש $|N(v)| < |c| - 1$.

הוכחת טענה ראשית: ע״פ הגדרת האלגוריתם מתקיים כי האלג׳ מחזיר $true \Leftrightarrow$ קיים $c \in C$ כך שקיים $v \in c$ כך ש $|N(v)| < |c| - 1$. מכאן, ע״פ טענת העזר מתקיים כי האלג׳ מחזיר $true \Leftrightarrow G$ בעל תת גרף מושרה P_3.

הוכחת טענת עזר: $\Leftarrow$ נניח כי G בעל תת גרף מושרה P_3. אזי קיימים $v, u, w \in V$ כך ש $(v, u), (u, w) \in E$ וגם $(v, w) \notin E$. נסמן $c \in C$ כך ש $v, u, w \in c$ ונבחין כי $|N(v)| < |c| - 1$. מכאן שקיים $c \in C$ כך שקיים $v \in c$ כך ש $|N(v)| < |c| - 1$.
$\Rightarrow$ נניח כי קיים $c \in C$ כך שקיים $v \in c$ כך ש $|N(v)| < |c| - 1$. נחלק למקרים:
I. $|c| = 1$: מכאן שלכל $u \in c$ מתקיים $|N(u)| = 0 = 1 - 1 = |c| - 1$ בסתירה להנחה.
II. $|c| = 2$: מכאן שלכל $u \in c$ מתקיים $|N(u)| = 1 = 2 - 1 = |c| - 1$ בסתירה להנחה.
III. $|c| \geq 3$: $|N(v)| < |c| - 1$ אזי קיים $w \in c$ כך ש $(v, w) \notin E$.
מכיוון ש $v, w \in c$ וגם $|c| \geq 3$ נקבל כי G מכיל מסלול $(p_1 = v), p_2, \ldots, (p_k = w)$ כך ש $k \geq 3$.
אם $k = 3$ אזי p_1, p_2, p_3 הינו תת גרף מושרה P_3 של G וסיימנו. אחרת $k > 3$ ונחלק למקרים:
III.I. קיים p_i עבור $2 \leq i \leq k - 1$ כך ש $(p_1, p_i), (p_i, p_k) \in E$. אזי p_1, p_i, p_k הינו תת גרף מושרה P_3 של G וסיימנו.
III.II. לכל p_i עבור $2 \leq i \leq k - 1$ מתקיים כי $(p_1, p_i) \notin E$ או $(p_i, p_k) \notin E$.
נסתכל על p_j בעל j מקסימלי עבור $2 \leq j \leq k - 2$ כך ש $(p_1, p_j) \in E$, ומכאן ש p_1, p_j, p_{j+1} הינו תת גרף מושרה P_3 של G וסיימנו.

זמן ריצה

1. $O(n + m)$.
2. $O(n)$.
3. $O(1)$.

סה״כ זמן ריצת האלגוריתם הינו $O(n + m)$.

שאלה 4

לכל $u, v \in V$ נגדיר $P(u,v)$: קבוצת הקודקודים הבונים את המסלול הקצר ביותר מ-u ל-v (במידה וקיימות מספר קבוצות כאלה, נבחר אחת מהן באופן אקראי).

תיאור האלגוריתם

1. $G' \leftarrow (V, E'): E' = \{\{u,v\}: \{v,u\} \in E\}, \qquad \hat{d} \leftarrow Matrix[n \times n]$

2. $for\ u, v \in V: \hat{d}(u,v) \leftarrow \infty$

3. $X \leftarrow Randomized\ Hitting\ Set(n = n, k = n^2, R = n^\epsilon)$

4. $for\ x \in X:$

 4.1 $D \leftarrow BFS(G, x)$

 4.2 $D' \leftarrow BFS(G', x)$

 4.3 $for\ u, v \in V:$

  4.3.1 $\hat{d}(u,v) \leftarrow \min\{\hat{d}(u,v),\ D'(u) + D(v)\}$

5. $return\ \hat{d}$

נשים לב :

- האלגוריתם הרנדומלי לחישוב $Hitting\ Set$ אינו דורש את תתי הקבוצות הדורשות פגיעה, נסמן תתי קבוצות אלה ב $P_\epsilon = \{P(u,v): |P(u,v)| \geq n^\epsilon\}$.
- בשורות 4.1, 4.2 מריצים סריקת BFS ומחזירים את וקטור המרחקים, כאשר D הינו וקטור המרחקים מ x לכל $v \in V$ ו D' הינו וקטור המרחקים מכל $v \in V$ ל x.

הוכחת נכונות

יהיו $u, v \in V$ כך ש $|P(u,v)| \geq n^\epsilon$, כלומר המסלול הקצר ביותר מ-u ל-v הוא באורך של לפחות n^ϵ.

לפי הגדרה $P(u,v) \in P_\epsilon$.

כפי שראינו בכיתה, האלגוריתם הרנדומלי לחישוב $Hitting\ Set$ מצליח בהסתברות גבוהה למצוא X נכון.

נניח שהאלגוריתם עבד ומצא X נכון, לכן קיים $x \in X$ כלשהו המקיים $x \in P(u,v)$.

<u>אבחנה</u> : תת מסלול של מסלול קצר ביותר הוא גם מסלול קצר ביותר.

<u>אבחנה</u> : $\forall a, b \in V: d_G(a,b) = d_{G'}(b,a)$ (מאחר ש G' הוא הגרף ההופכי ל-G).

ע"פ אבחנות אלה ומנכונות BFS נקבל כי $d(u,v) = d(u,x) + d(x,v)$.

לכן, האלגוריתם יחזיר בהסתברות גבוהה $\hat{d}(u,v) = d(u,v)$ כנדרש.

זמן ריצה

1. $O(n^2)$.
2. $O(n^2)$.
3. $O\left(\frac{2n}{R}\ln n\right) = O(n^{1-\epsilon} \cdot \ln n)$.
4. הלולאה הראשונה רצה $O(n^{1-\epsilon} \cdot \ln n)$ איטרציות, כאשר כל איטרציה לוקחת $O(n^2)$ ובסה״כ $O(n^{1-\epsilon} \cdot \ln n) \times \left(O(n^2) + O(n^2) + O(n^2)\right) = O(n^{3-\epsilon} \cdot \ln n)$.
5. $O(1)$.

בסה״כ זמן ריצה האלגוריתם הינו $\boldsymbol{O(n^{3-\epsilon} \cdot \ln n)}$ כנדרש.