Ben-Gurion University of the Negev

SOLUTION
By
Avi Ferdman

עבודת בית 1: תכנון אלגוריתמים 2020

תאריך הגשה: 10.11.2019, 12:00 בצהריים, תאים מספר 95,96 בקומת כניסה של בניין 37. כמו כן, יש להגיש עותק של העבודה במערכת ההגשה.

מתרגל אחראי: תומר סידי.

הוראות כלליות:

- כל עוד לא נאמר אחרת, כאשר הנכם מתבקשים לתאר אלגוריתם יש לספק את הבאות:
  1. תיאור מילולי של האלגוריתם.
  2. הוכחת נכונות.
  3. ניתוח זמן ריצה (**לא כולל** זמן הריצה של הקופסה השחורה).
- אלגוריתם עם זמן ריצה אקספוננציאלי לא נחשב יעיל ולכן בדרך כלל לא יתקבל.
- פתרון יש לכתוב רק בדף התשובות הנלווה לעבודה.

<u>תזכורת</u>:

רדוקציה הינה פתרון בעיה אחת בעזרת בעיה אחרת. באופן פורמלי (מהתרגול):

הגדרה פורמאלית ל- "רדוקציה":

תהיינה A ו-B זוג בעיות נתונות. רדוקציה מ- A לבעיה B היא זוג פונקציות f,g כך ש:

- f היא פונקצית המרת הקלט, המעבירה מופע של בעיה A למופע של בעיה B.
- g היא פונקצית המרת הפלט, המעבירה פתרון של בעיה B לפתרון של בעיה A.
- עבור מופע a לבעיה A, אם $B(f(a))$ הוא פתרון עבור המופע $f(a)$ תחת בעיה B אזי $g(B(f(a)))$ הוא פתרון למופע a תחת בעיה A. (הגדרת נכונות)

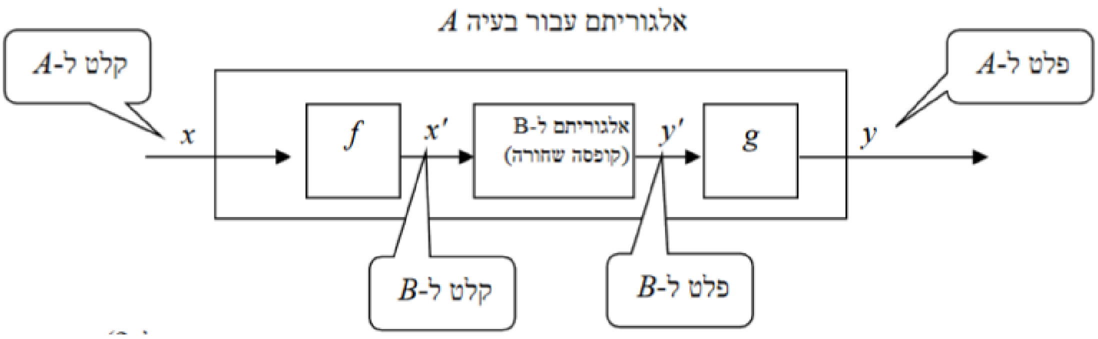
כדי להוכיח את נכונות הרדוקציה, יש להוכיח שהאלגוריתם הבא פותר את הבעיה A:

1. עבור מופע a לבעיה A, חשב את $f(a)$.
2. עבור המופע $f(a)$ לבעיה B, חשב את הפיתרון b.
3. חשב את $g(b)$ להיות הפתרון של A.

כאשר f הינה תרגום הקלט, g הינה תרגום הפלט והאלגוריתם לבעיה B הינו ה"קופסא השחורה".

לטובת העבודה, כשאנו מבקשים רדוקציה מבעיה A לבעיה B, יש לתת פתרון לבעיה A באמצעות קופסא שחורה של בעיה B.

<u>הערה</u>: ניתן להניח כי ממיר הפלט מכיר גם את הקלט המקורי a.



שאלה 1 (25 נקודות)

יהי $G=(E,V)$ ו- $(s,\ u_1,\ u_2,\ t)$ סדרה של קוד' שונים בגרף כך ש- $s,\ u_1,\ u_2,\ t\ \epsilon\ V$.

שלב תרגום הקלט:

נסמן:

$$V_1=\{v^1\mid v^1\ \epsilon\ V\ \wedge\ v^1\ \neq\ u_1\wedge\ v^1\ \neq\ u_2\}$$

$$E_1=\{(v^1,u^1)\ \mid v^1,u^1\ \epsilon\ V_1\ \wedge(v^1,u^1)\ \epsilon\ \mathrm{E}\}\cup\{(v^1,u_1)\mid v^1\ \epsilon\ V_1\ \wedge(v^1,u_1)\ \epsilon\ \mathrm{E}\}$$

$$V_2=\{v^2\mid v^2\ \epsilon\ V\ \wedge\ v^1\ \neq\ u_2\}$$

$$E_2=\{(v^2,u^2)\ \mid v^2,u^2\ \epsilon\ V_2\ \wedge(v^2,u^2)\ \epsilon\ \mathrm{E}\}\cup\{(v^2,u_2)\mid v^2\ \epsilon\ V_2\ \wedge(v^2,u_2)\ \epsilon\ \mathrm{E}\}$$

$$V_3=\{v^3\mid v^3\ \epsilon\ V\ \wedge\ v^3\ \neq\ u_1\}$$

$$E_3=\{(v^3,u^3)\ \mid v^3,u^3\ \epsilon\ V_3\ \wedge(v^3,u^3)\ \epsilon\ \mathrm{E}\}$$

נבנה גרף חדש, נגדיר: $\mathrm{G}'=(\mathrm{V}',\mathrm{E}')\quad V'=V_1\cup V_2\cup V_3\quad \mathrm{E}'=E_1\cup\ E_2\cup\ E_3$

שלב תרגום הפלט:

הפלט נותר ללא שינוי. כלומר הפלט של הקופסא השחורה יהיה אורך המסלול הקצר ביותר החוקי המבוקש.

הוכחה:

בהנתן גרף $G=(E,V)$ ו- $(s,\ u_1,\ u_2,\ t)$ סדרה של קוד' שונים בגרף כך ש- $s,\ u_1,\ u_2,\ t\ \epsilon\ V$.

טענה ראשית:

לכל סדרה של קוד' $s,\ u_1,\ u_2,\ t\ \epsilon\ V$ אורך המסלול הקצר ביותר מ- s ל- t שעובר דרך הקוד $u_1,\ u_2$ לפי סדר כאשר כל אחד מהם מופיע פעם אחת בלבד, שווה לאורך המסלול $d(s^1,t^3)$ בגרף G'.

טענת עזר:

קיים ב-G מסלול מאורך l (∞) בין s ל- t שעובר דרך הקוד $u_1,\ u_2$ לפי סדר כאשר כל אחד מהם מופיע פעם אחת בלבד אם"ם קיים ב-G' מסלול מאורך l (∞) בין s^1 ל- t^3.

הוכחת הטענה הראשית ע"י טענת העזר:

יהי l אורך המסלול המינימאלי ב-G מ- s ל- t שעובר דרך הקוד $u_1,\ u_2$ לפי הסדר כאשר כל אחד מהם מופיע פעם אחת בלבד, מטענת העזר קיים ב-G' מסלול מאורך l מs^1 ל- t^3. נניח בשלילה שב-G' יש מסלול מs^1 ל- t^3 באורך k כך ש: $k<l$. מטענת העזר קיים ב-G מסלול בין s ל- t שעובר דרך הקוד $u_1,\ u_2$ לפי סדר כאשר כל אחד מהם מופיע פעם אחת בלבד באורך k, סתירה למינימאליות l.

הוכחת טענת העזר:

$\Leftarrow$ מקרה א: נניח כי קיים בG מסלול מ- s ל- t שעובר דרך הקוד $u_1,\ u_2$ לפי הסדר בגודל l צ"ל קיים ב'G מסלול מאורך l בין s^1 ל- t^3.יהי P מסלול מ-s ל-t שעובר דרך הקוד $u_1,\ u_2$ לפי הסדר בגודל l נסמנו:

$$P=(s,v_1,v_{2,},\dots,v_i,u_1,v_{i+1},\dots,v_j,u_2,v_{j+1},\dots,t\ )$$

נתבונן ב-3 חלקים של המסלול:

$$P_1 = (\text{s}, v_1, v_{2,} \ldots, v_i)$$

$$P_2 = (u_1, v_{i+1}, \ldots, v_j)$$

$$P_3 = (u_2, v_{j+1}, \ldots, t\ )$$

כעת נרכיב באמצעות שלושת הקבוצות הללו מסלול חוקי באורך l ב-'G. נתבונן ב-P_1, נשים לב כי: $\forall v\epsilon P_1: v\epsilon V_1$. זאת משום שכל הקוד' בגרף המקורי שייכים לV_1 פרט לu_1 ו-u_2 אבל קוד' אלה אינם נמצאים בP_1, מכיוון שאילו היו נמצאים בP_1 הם היו מופיעים יותר מפעם אחת בP כיוון שאחד מהם נמצא בP_2 והשני בP_3 וזו סתירה. נשים לב כי כל קשת במסלול P_1 נמצאת גם בE_1, זאת מכיוון שקבוצת הקשתות E_1 אינה משנה את קבוצת הקשתות המקורית פרט לקשתות אחד מהקוד' שאותם היא מחברת הם u_1 או u_2. אבל כמו שראינו הקוד הללו אינם שייכים לP_1. לכן ניתן ליצור ב'G את המסלול: $P_1' = (s^1, v_1^1, v_2^1, \ldots, v_i^1)$. נתבונן ב-$P_2$, נשים לב כי: $\forall v\epsilon P_2: v\epsilon V_2$. זאת משום שכל הקוד' בגרף המקורי שייכים לV_1 פרט לu_2 אבל קוד' זה אינו נמצאים בP_2, מכיוון שאילו היה נמצא בP_2 היה מופיע יותר מפעם אחת בP מכיוון שהוא שייך לP_2, וזו סתירה. נשים לב כי כל קשת במסלול P_2 נמצאת גם בE_2, זאת מכיוון שקבוצת הקשתות E_2 אינה משנה את קבוצת הקשתות המקורית פרט לקשתות שאחד הקוד שלהם הוא u_2 אבל ראינו שהקוד' הזה אינו שייך לP_2 לכן ניתן ליצור ב'G את המסלול:

$P_2' = (u_1, v_{i+1}^2, \ldots, v_j^2)$ באופן זהה ליצירת המסלול החוקי P_2' ב-'G ניתן ליצור את המסלול: $P_3' = (u_2, v_{j+1}^3, \ldots, t^3)$ מכיוון שבמסלול המקורי P קיימת הקשת (u_i, u_1) הקשת$(v_i^1, u_1)\epsilon E_1$ לפי הגדרה. משיקולים דומים: $(v_j^2, u_2)\epsilon E_2$. לכן ב-'G ניתן ליצור את המסלול: $\text{P}' = P_1' \cup P_2' \cup P_3'$ מהבנייה של P_1' דרך P_1 נובע: $|P_1'| = |P_1|$. באופן דומה: $|P_2'| = |P_2|$ וגם $|P_3'| = |P_3|$ולכן:$|P| = |\text{P}'| = \text{l} + 1$ ולכן נובע כי אורך המסלול 'P ב-'G הוא l.

מקרה ב: נניח כי לא קיים בG מסלול מ- s ל- t שעובר דרך הקוד $u_2,\ u_1$ לפי הסדר צ"ל לא קיים ב'G מסלול בין s^1 ל- t^3. נניח בשלילה שכן קיים מסלול ב'G בין s^1 ל- t^3. אם כך נוכל לבנות מסלול ב-G מ- s ל- t שעובר דרך הקוד $u_2,\ u_1$ לפי הסדר בדיוק כמו שיצרנו מסלול כזה תחת ההנחה בהוכחת הכיוון השני ונקבל סתירה.

<u>$\Longrightarrow$</u> מקרה א: נניח כי קיים ב'G מסלול באורך l מ-s^1 ל-t^3 צ"ל קיים ב-G מסלול מאורך l בין s ל- t שעובר דרך הקוד $u_2,\ u_1$ לפי סדר כאשר כל אחד מהם מופיע פעם אחת.יהי

$P' = (s^1, v_1^1, v_2^1, \ldots, v_i^1, u_1, v_{i+1}^2, \ldots, v_j^2, u_2, v_{j+1}^3, \ldots, t^3)$ מסלול כזה ב-G'. לפי הגדרת G', $\forall i: 1 \le i \le 3$ מתקיים אם:

$(v^i, u^i)\epsilon E'$ אז: $(v, u)\epsilon E$ בנוסף מתקיים אם: $(v^1, u_1)\epsilon E'$ אז: $(v, u_1)\epsilon E$ מתקיים אם: $(v^2, u_2)\epsilon E'$ אז: $(v, u_2)\epsilon E$

לכן ניתן לבנות מסלול חוקי בG באופן הבא:

$P = (\text{s}, v_1, v_{2,} \ldots, v_i, u_1, v_{i+1}, \ldots, v_j, u_2, v_{j+1}, \ldots, t\ )$ זהו מסלול חוקי כי בין כל שני קוד' קיימת קשת, והמסלול הזה הוא באורך l מכיוון שהתחלנו עם מסלול בגודל l ולכל קוד' במסלול 'P התאמנו קודקוד בP. כעת נותר להראות שהקוד' u_1, u_2 מופיעים פעם אחת בלבד בP. נניח בשלילה שקוד' u_1 מופיע פעמיים לפחות בP. ז"א שב-'P קיימות לפחות שתי קשתות: (x_1, u_1), (x_2, u_1) אבל זו סתירה לחוקיותו של המסלול P' כי u_1 מופיע בו פעם אחת בלבד. כעת נראה שקוד' u_1 מופיע לכל היותר פעם אחת, וזו מכיוון שבמסלול 'P מופיעה הקשת (v_i^1, u_1) ולכן במסלול P מופיעה הקשת:(v_i, u_1). באופן דומה ניתן להוכיח שהקוד' u_2 מופיע בדיוק פעם אחת בP.

מקרה ב: נניח כי לא קיים ב'G מסלול באורך l מ-s^1 ל-t^3 צ"ל לא קיים ב-G מסלול מאורך l בין s ל- t שעובר דרך הקוד $u_2,\ u_1$ לפי סדר כאשר כל אחד מהם מופיע פעם אחת. נניח בשלילה שכן קיים ב-G מסלול מאורך l בין s ל- t שעובר דרך הקוד $u_2,\ u_1$ לפי סדר אז נוכל לבנות ב'G מסלול מ-s^1 ל-t^3 באותו אופן כמו שבנינו מסלול כזה בהוכחת הכיוון השני ונגיע לסתירה.

<u>תיאור האלגוריתם:</u>

תרגם את הקלט.

הפעל את הקופסא השחורה על $d(s^1, t^3)$

החזר את הפלט של הקופסא השחורה.

<u>ניתוח זמן ריצה:</u>

תרגום הפלט:

משכפלים את כל הקוד' 3 פעמים (פרט ל2 הקוד u_1, u_2): $O(|V|)$

משכפלים את כל הקשתות 3 פעמים (פרט לקשתות המחוברות לקוד' u_1, u_2): $O(|E|)$

את הקשתות שאחת מהקוד' שלהן הם: u_2 או u_1 מעתיקים לגרף החדש פעם אחת: $O\left(N(u_1)+N(u_2)\right)=O(|V|)$

תרגום הפלט:

לא שינוי: $O(1)$

לכן סה"כ יעילות האלגוריתם ל כולל זמן הריצה של הקופסא השחורה היא: $O(|V|+|E|)$

<u>שאלה 2 (25 נקודות)</u>

יהי $G=(V,E)$ גרף לא מכוון ו-2 קבוצות קוד' $U,M \subset V$ כך ש: $U \cap M = \emptyset$ וגם: $U \cup M = V$

<u>שלב תרגום הקלט:</u>

נסמן:

$$E' = \{(v,u) \mid u \in U \ \wedge v \in M \ \wedge (v,u) \in E\}$$

נבנה גרף חדש: $G' = (V,E')$

<u>שלב תרגום הפלט:</u>

נחזיר את קב' כיסוי בקודקודים בגודל מינימאלי של הגרף G'.

<u>הוכחה:</u>

בהנתן $G=(V,E)$ גרף לא מכוון ו-2 קבוצות קוד' $U,M \subset V$ כך ש: $U \cap M = \emptyset$ וגם: $U \cup M = V$, קבוצת חלוקה מינימאלית שווה לקבוצת כיסוי בקוד' בגודל מינימאלי בגרף G'.

<u>טענת עזר מספר 1:</u>

בגרף G קבוצה $L \subset V$ מהווה חלוקה אם"ם הקבוצה L מהווה כיסוי בגרף G'.

<u>הוכחת הטענה הראשית:</u>

יהי $G=(V,E)$ גרף לא מכוון ו-2 קבוצות קוד' $U,M \subset V$ כך ש: $U \cap M = \emptyset$ וגם: $U \cup M = V$ ותהי $L \subset V$ קבוצת חלוקה מינימאלית של G. לפי טענת עזר מס' 1 נובע כי L היא כיסוי בקוד' של הגרף G'. כעת נניח בשלילה כי קיים כיסוי אחר של G' נסמנו K כך ש: $|K| < |L|$. לכן לפי טענת עזר מספר 1 הקבוצה K מהווה חלוקה של הגרף G, אבל זו סתירה למינימאליות של L כי: $|K| < |L|$.

<u>הוכחת טענת העזר:</u>

<u>ניעזר בטענת עזר נוספת: טענת עזר מספר 1.1:</u>

יהי $G=(V,E)$ גרף לא מכוון ותהי $V' \subseteq V$ אם $E = \{(u,v) \mid v \in V'\}$ אזי V' מהווה כיסוי של G.

$\Leftarrow$ יהי $G=(V,E)$ גרף לא מכוון ו-2 קבוצות קוד' $U,M \subset V$ כך ש: $U \cap M = \emptyset$ וגם: $U \cup M = V$ ותהי L חלוקה של הגרף G צ"ל L כיסוי של G'. לפי הגדרת $G' = (V,E')$ ולפי הגדרת קבוצת החלוקה נובע כי $E' = \{(u,v) \mid v \in L\}$. נניח בשלילה ש: $E' \neq \{(u,v) \mid v \in L\}$ אז נובע שקיימת קשת $(v,u) \in E'$ כך ש: $v,u \notin L$ אם כך $v,u \in M$ או $v,u \in U$ *ובכל מקרה לפי הגדרת* E' : *$(v,u) \notin E'$ סתירה*. לכן לפי טענת עזר 1.1 L מהווה כיסוי של G'.

$\Rightarrow$ תהי קבוצה L כיסוי ב- G' צ"ל L חלוקה ב-G. נניח בשלילה ש-L אינה חלוקה ב-G, אז קיימים 2 קוד $u \in U\setminus L,\ v \in M\setminus L$ שנמצאים באותו רכיב קשירות. לכן קיימים קוד' : $u' \in U\setminus L,\ v' \in M\setminus L$ לפי בניית G' : $(v',u') \in E'$ לכן לפי הגדרת הכיסוי נובע:

$u',v' \in L$ או $u' \in L$ או $v' \in L$ ובכל מקרה זו סתירה כי הנחנו ש: $u' \in U\setminus L,\ v' \in M\setminus L$.

<u>הוכחת טענת העזר 1.1:</u>

יהי $G = (V, E)$ גרף לא מכוון ותהי $V' \subseteq V$ אם $E = \{(u, v) \mid v \epsilon V'\}$ צ"ל 'V מהווה כיסוי של G. נניח בשלילה ש- 'V אינה מהווה כיסוי בG , כלומר קיימת קשת $(u, v) \in E$ כך ש: $u, v \notin V'$ אבל זו סתירה להנחה.

<u>תיאור האלגוריתם:</u>

תרגם את הקלט.

הפעל את הקופסא השחורה על G'

החזר את הפלט של הקופסא השחורה. (קבוצת כיסוי מינימאלית ב'G).

<u>ניתוח זמן ריצה:</u>

<u>תרגום הקלט:</u>

משכפלים כל קוד' בגרף: O(|V|)

עוברים על כל הקשתות שבגרף (בכדי להבחין בין קשתות שמחברות את הקבוצות U ו-M): O(|E|)

<u>תרגום הקלט:</u>

ללא שינוי: $O(1)$

סה"כ זמן הריצה של האלגוריתם לא כולל זמן הריצה של הקופסא השחורה: $O(|V| + |E|)$

שאלה 3 (25 נקודות)

יהיו $x_1, x_2, \dots\ x_n$ מספרים רציונליים שונים זה מזה.

שלב תרגום הקלט:

נמיין את $x_1, x_2, \dots\ x_n$ כך שx_1 הוא הקטן ביותר. נבנה גרף $G = (V, E)$ מכוון וממושקל באופן הבא:

$$V = \{v_1, v_2, \dots\ v_n, v_{n+1}\}$$

כל קוד' למעט הקודקוד האחרון, ייצג נקודת קלט. נגדיר $v_i =\ x_i : \forall i{:}\ 1 \le i \le n$

$$E = \{(v_i, v_j) \mid v_i < v_j\} \cup \{(v_i, v_{n+1}) | 1 \le i \le n\}$$

נגדיר פונקציית משקל:

תהי קשת $(v_i, v_{j+1}) \in E$ נגדיר בהתאם: $M = (v_i, v_{i+1}, \dots, v_j)$ נסמן את החציון של הקבוצה M ב-c. נגדיר:

$$cost(M) = \sum_{l=i}^{j} |x_l - c|$$

$$W{:}\ (v_i, v_{j+1}) \longrightarrow Cost(M)$$

שלב תרגום הפלט:

הפלט הוא מסלול בגרף, נסמנו: $P = (u_1, u_2, \dots, u_{k+1})$. נאתחל קבוצה ריקה שתכיל בעתיד את החציונים. לכל קשת (u_i, u_{i+1}) במסלול, נחשב את החציון בקטע חצי פתוח: $[u_i, u_{i+1})$. נוסיף חציון זה לקבוצת החציונים.

תיאור האלגוריתם:

נמיין את $x_1, x_2, \dots\ x_n$ כך שx_1 הוא הקטן ביותר. נבנה גרף $G = (V, E)$ מכוון וממושקל באופן הנ"ל. נשתמש בקופסא השחורה עם הפרמטרים: (v_1, v_{n+1}, k). נמיר את הפלט.

הוכחה:

בהנתן $x_1, x_2, \dots\ x_n$ נקודות קלט, ומספר טבעי k.

טענה ראשית:

האלגוריתם מחזיר קבוצה M של k מספרים רציונאליים $M = \{m_1, \dots, m_k\}$, שעבורם העלות ביחס לנקודות הקלט המוגדת באופן הבא: $cost(M) = \sum_{i=1}^{n} min_{j=1\dots k}\ \{|x_i - m_j|\}$ מיניאלית.

טענת עזר:

לכל קבוצה $M_k = \{m_1, m_2, \dots, m_k\}$ קיים מסלול P באורך k שמשקלו W(P) כך שמתקיים: $W(P) \le cost(M_k)$ בגרף G באורך k.

טענת עזר 2:

האלגוריתם החזיר קבוצה M מגודל k $\Longleftarrow$ קיים מסלול P כך שמתקיים $W(P) = cost(M)$

הוכחת הטענה הראשית בעזרת טענת העזר:

תהי M_1 הקבוצה שהאלגוריתם החזיר. לפי טענת עזר 2: האלגוריתם החזיר מסלול P באורך k כך שמתקיים

$$W(P) = cost(M_1)$$

נניח בשלילה שהאלגוריתם אינו נכון וקיימת קבוצה M_2 בגודל k שעבורה מתקיים $cost(M_2) < cost(M_1)$. אם כך לפי טענת העזר קיים מסלול P_2 כך שמתקיים: $W(P_2) \leq cost(M_2) < cost(M_1) = \text{W(P)}$ ולכן בפרט נובע $\text{W}(P_2) < \text{W(P)}$ אבל זו סתירה כי מנכונות הקופסא השחורה היא מחזירה את המסלול בעל המשקל הקטן ביותר מאורך k ולכן לא יכול להיות שקיים מסלול אחר מאורך k עם משקל קטן יותר כי אחרת הקופסא השחורה היתה מחזירה אותו. ולכן הקבוצה שהאלגוריתם החזיר היא בעלת עלות מינימאלית.

<u>הוכחת טענת העזר 1:</u>

תהי $M = \{m_1, m_2, \ldots, m_k\}$ צ"ל קיים מסלול P ב-G מגודל k שעבורו: $W(P) \leq cost(M)$ניצור קבוצה חדשה נסמנה 'M כך שמתקיים $cost(M') \leq cost(M)$ ונראה כי $W(P) \leq cost(M')$ואז בפרט יתקיים אי השיוויון שצריך להוכיח: $W(P) \leq cost(M)$. יהיו $I_1, I_2, \ldots, I_k$ תתי קבוצות של הקבוצה $\{x_1, x_2, \ldots, x_n\}$ כך שלכל I_i בחישוב ה$cost(M)$ המרחק של כל האיברים ששייכים אליה מחושב מאותה "נקודת מרכז" $m_i \epsilon \text{M}$. נסדר את הקבוצות I לפי סדר האיבר הקטן ביותר בכל קבוצה וניצור את הקבוצה 'M בעזרת 3 הבחנות: (כלומר I_i תהיה בסדר לפני I_j אם האיבר הקטן ביותר בI_i קטן מהאיבר הקטן ביותר בI_j)

- הקבוצות $I_1, I_2, \ldots, I_k$ זרות בזוגות. מכיוון שאילו הן אינן זרות בזוגות בM יש לפחות איבר אחד x ששייך לפחות ל2 קבוצות נסמנן I_i, I_j ואז נשמיט את x מאחת הקבוצות הללו ונקבל קבוצה חדשה 'M שעבורה מתקיים $cost(M') \leq cost(M)$
- $\forall i < j \leq k$ מתקיים $\forall e_i, e_j | e_i \epsilon I_i, e_j \epsilon I_j$ מתקיים: $e_i < e_j$. מכיוון שאלמלא כן יש לפחות שני איברים $e_i, e_j | e_i \epsilon I_i, e_j \epsilon I_j$ כך ש: $e_j < e_i$ ואז נחליף את e_j, e_i נשים כל אחד מהם בקבוצה של השני ונקבל קבוצה חדשה 'M שעבורה מתקיים $cost(M') \leq cost(M)$
- כל נקודה $m_i \in M$ היא חציון של הקבוצה I_i המתאימה לה (כלומר m_i היא הנקודה שאליה נמדדים המרחקים של הנקודות ששייכות אך ורק לI_i בחישוב $cost(M)$) מכיוון שאחרת אם היא לא החציון ונסמן את החציון ב-c מהטענה הידועה אם c חציון של קבוצה $\{x_1, x_2, \ldots, x_n\}$ מתקיים:

$$\forall m \colon \sum_{i=1}^{m} |x_i - c| = min_{d \epsilon R} \sum_{i=1}^{m} |x_i - d|$$

נוכל להחליף את m_i ב-c וע"י כך **לא הגדלנו** את $cost(M)$ ולכן נקבל קבוצה חדשה של מרכזים 'M שעבורה מתקיים $cost(M') \leq cost(M)$ (נוכל לעשות את זה עבור כל $m_i \epsilon M$ שאינו חציון של תת קבוצה מתאימה I_i).

לכן לאחר התהליך הזה בהכרח נקבל קבוצה חדשה נסמנה 'M כך שמתקיים $cost(M') \leq cost(M)$. כעת נותר להראות כי קיים P כך ש- $W(P) \leq cost(M')$. יהיו $I_1', I_2', \ldots, I_k'$ תתי קבוצות של הקבוצה $\{x_1, x_2, \ldots, x_n\}$ כך שלכל I_i בחישוב ה$cost(M')$ המרחק של כל האיברים ששייכים אליה מחושב מאותה "נקודת מרכז" $m_i' \epsilon \text{M}'$. מאופן הבניה של 'M נוכל לסדר את הקבוצות $I_1, I_2, \ldots, I_k$ לפי סדר האיבר הקטן ביותר נסמן:

$$((v_1, v_2, \ldots, v_i), (v_{i+1}, v_{i+2}, \ldots, v_j), \ldots, (v_{m+1}, v_{m+2}, \ldots, v_l), (v_{l+1}, v_{l+2}, \ldots, v_n))$$

מכיוון שלפי הגדרת הגרף G קיימת קשת בין כל איבר ראשון של קבוצה כלשהי לבין כל שאר האיברים הגדולים ממנו **ובפרט לאיבר הראשון של הקבוצה הבאה אחריה בסדר**, ניתן להרכיב את המסלול: $\text{P} = (v_1, v_{i+1}, \ldots, v_{m+1}, v_{l+1}, v_{n+1})$ שזהו מסלול חוקי בגרף לפי הגדרת פונקציית המשקל בגרף ולפי אופן בניית 'M המשקל הכולל של המסלול הוא בדיוק $cost(M')$.

ולכן קיבלנו מסלול P ומתקיים: $W(P) = cost(M') \leq cost(M)$

<u>הוכחת טענת עזר 2:</u>

$\Leftarrow$ תהי קבוצה $M=\{m_1,m_2,\dots,m_k\}$ מגודל k הפלט של האלגוריתם צ"ל קיים מסלול P ב-G מגודל k כך ש: $cost(M)=W(P)$. לפי הגדרת ממיר הקלט הקבוצה M היא החציונים של הקטעים: $[u_i,u_{i+1})$ לכל: $1\le i\le k$ לכן לפי הגדרת ממיר הפלט קיימות הקשתות: (u_i,u_{i+1}) לכל: $1\le i\le k$ סדרת הקשתות האלה יוצרות מסלול P ולפי הגדרת פונקציית המשקל W נובע $\mathrm{W}(P)=cost(M)$

<u>ניתוח זמן ריצה:</u>

מיון n הקלטים: $O(n\log n)$

שלב תרגום הקלט:

יצירת N+1 קוד': $O(n)$

יצירת קשת בין כל 2 קוד ($O(n^2)$) ולכל קשת חישוב פונקציית המשקל שלה.

חישוב פונקציית המשקל מתבצע ע"י מעבר על כל הקוד שנמצאים בין הקוד שיוצאת ממנו הקשת לבין הקוד

שהקשת נכנסת אליו.

הקשת הארוכה ביותר (קיימת אחת כזו) מכילה בתוכה n קוד שצריך לחשב עבורם את העלות,

הקשת השניה בגודלה (קיימות שתיים כאלה) מכילות בתוכן n-1 קוד שצריך לחשב עבורתם את העלות,

לכן נוצר זמן חישוב:

$$\sum_{k=1}^{n}(n-k+1)k=(n+1)\sum_{k=1}^{n}k-\sum_{k=1}^{n}k^2=(n+1)[0.5n(n+1)]=O(n^3)$$

תרגום הקלט:

לכל קשת במסלול הפלט מחשבים את החציון שבקטע בין הקוד, שהקשת יוצאת ממנו **כולל** לבין הקוד' שהקשת נכנסת אליו **לא כולל**. סך הכל עוברים על כל הקוד' מספר קבוע של פעמים תלוי במימוש: O(n)

לכן בסך הכל זמן הריצה הכללי של האלגוריתם הוא: $O(n^3)$

<u>שאלה 4 (25 נקודות)</u>

סעיף א

עבור כל קוד' x ϵ V:

עבור כל קוד $y \epsilon N(x)$

עבור כל קוד $z \epsilon N(y)$

אם יש קשת בין z ל- x

תחזיר true

תחזיר false

(N – מייצג את קבוצת השכנים)

סעיף ב

יהיו $v_i, v_j \epsilon V$

⇐ נניח שיש מסלול באורך 2 בין v_j ל v_i ב G_A צ"ל : יש קשת בין v_j ל v_i ב G_C נרצה להראות ש:
$C_{i,j} = 1$. מהגדרת $C_{i,j}$ מספיק למצוא שכן משותף אחד $v_k \epsilon V$ ל v_j ו v_i כדי שתהיה קשת בין v_j ל v_i ב G_C כלומר $C_{i,j} = 1$.
מההנחה יש מסלול באורך 2 בין v_j ל v_i ב G_A לכן קיים $V_k \epsilon V$ כך שמתקיים: $(v_i, v_k), (v_k, v_j) \epsilon E$ לכן בפעולת הכפלת המטריצה בעצמה ולפי הגדרת $C_{i,j}$ נקבל ש ש $C_{i,j} = 1$ לכן יש קשת בין v_j ל v_i ב G_C כנדרש.

⇒ נניח שיש קשת בין v_j ל v_i ב G_C צ"ל: יש מסלול באורך 2 בין v_j ל v_i ב G_A מההנחה שיש קשת בין v_j ל v_i ב G_C ומהגדרת $C_{i,j}$ נקבל שקיים $v_k \epsilon V$ כך ש $A_{i,k} \wedge A_{k,j} = 1$ כלומר קיים להם שכן משותף ב- G_A לכן $(v_i, v_k), (v_k, v_j) \epsilon E$ לכן יש מסלול באורך 2 בין v_j ל v_i ב G_A כנדרש.

סעיף ג

יהי $G = (V, E)$ גרף לא מכוון עם n קודקודים.

<u>שלב תרגום הקלט:</u>

נסמן n=|V|. נבנה מטריצת שכנויות A בגודל n*n ונאתחל ערכים של 0 בתאים. לכל $1 \leq i, j \leq n$ אם $(v_i, v_j) \epsilon E$ נסמן $A_{i,j} =$
1

שלב תרגום הפלט:

נעבור על כל התאים בהם הערך הוא 1 (כלומר קיימת קשת) המתקבלים לאחר כפל מטריצת השכניות בעצמה ונבדוק לכל תא במיקום i,j שהערך בו 1 אם במטריצת השכנויות A המקורית הערך במיקום ה-i,j הוא גם כן 1 כלומר קיימת קשת בין 2 הקודקדים בגרף המקורי וגם מסלול באורך 2 (לפי הסעיף הקודם) כלומר קיים משולש. אם כן נחזיר "כן" (כלומר שיש מעגל פשוט באורך 3). אם לא נמצא תא כזה נחזיר "לא".

תיאור האלגוריתם :

נתרגם את הקלט למטריצת שכניות, נפעיל את הקופסא השחורה ונמיר את הפלט.

הוכחה :

בהנתן גרף $G = (V, E)$ לא מכוון עם n קודקודים

טענה ראשית:

קיים בגרף משולש אמ"ם קיימים 2 אינדקסים i, j כך ש $C_{i,j} = 1$ (C מטריצת השכוניות המתקבלת מכפל של A בעצמה) וגם $A_{i,j} = 1$.

טענה עזר:

יש מסלול באורך 2 בין v_j ל v_i ב G_A אמ"ם יש קשת בין v_j ל v_i ב G_C.

הוכחת הטענה הראשית ע"י טענת העזר:

$\Leftarrow$ נניח כי קיים בG משולש צ"ל קיימים 2 אינדקסים i, j כך ש $C_{i,j} = 1$ בעצמה) וגם $A_{i,j} = 1$. נסמן את המסלול הזה: $P = (v_i, v_j, v_l, v_i)$. לכן בפרט קיימת הקשת $(v_{i,}, v_j)$. לכן $A_{i,j} = 1$. בנוסף קיים מסלול באורך 2 מ-v_j ל-v_i ולכן לפי טענת העזר נובע כי: $C_{i,j} = 1$.

$\Rightarrow$ תהי A מטריצת שכניות המתקבלת ממעבר על כל הצלעות של G. תהי C מטריצת השכניות המתקבלת מכפל של A בעצמה יהיו 2 קודקדים v_i, v_j כך ש $C_{i,j} = 1$ וגם $\mathrm{A}_{i.j} = 1$. מכיוון ש: $C_{i,j} = 1$ לפי טענת העזר נקבל שיש מסלול באורך 2 בG_A בין v_i ל-v_j, נסמנו $P_1 = (v_i, v_l, v_j)$. בנוסף קיימת קשת בין v_j ל v_i ב G_A לכן קיים מסלול באורך 1 בין v_j ל- v_i נסמנו: $P_2 = (v_j, v_i)$.
כעת נבצע: $P_1 \cup P_2 = (v_i, v_l, v_j, v_i)$ ונקבל מסלול שהוא מעגל פשוט באורך 3 ב G_A כנדרש.

הוכחת הטענת העזר :

סעיף ב

ניתוח זמן ריצה :

בניית מטריצת שכניות A מגרף G (מעבר על הקשתות) - O(|E|)

בניית מטריצה חדשה C – $O(n^w)$

תרגום הפלט – מעבר כל תא ב C ובדיקה אם אותו תא ב A מכיל את הערך 1 – מעבר על כל התאים $O(n^2)$

סה"כ : $O(n^2) + O(n^w) + O(n^2) = O(|E|) + O(n^w) + O(n^2)$ => $O(n^2 + n^w)$

בהצלחה!