ממן 12

הפתרון נכתב על-ידי שלומי וקנין

<u>שאלה 1</u>

טענה1: הסרה של קשת e מעץ פורש מינימאלי T, תשאיר 2 תתי עצים מינימאליים אשר איחודם כולל את כל הצמתים בגרף.

הוכחה: ראשית נראה כי איחודם כולל את כל הצמתים: נניח כי קיימים n צמתים בגרף ו T הוא עץ פורש. לאחר הסרת הקשת e נקבל שני תתי עצים, a צמתים בתת עץ אחד, ו b צמתים בתת עץ השני, כך ש a+b=n, ולכן נוכל להגיע אל כל צומת דרך אחד משני העצים.

<u>נוכיח מינימאליות</u>: יהי w(T) העלות של העץ המינימאלי, ו w(A) ו w(B) העלויות של שני תתי העצים שהתקבלו בהתאמה. מכך ששני תתי העצים התקבלו מהסרת הקשת e, מתקיים w(T) = w(A)+w(B)+w(e)

נניח בשלילה וללא הגבלת הכלליות כי תת העץ A אינו מינימאלי, כלומר קיים סידור אחר של צמתי A כך שמתקבל תת עץ `A, המקיים w(A`)<w(A). מכאן, נוכל להחזיר את הקשת e ולקבל w(A`)+w(B)+w(e)<w(T) – בשלילה למינימאליות של T.

טענה 2: הוספת קשת בעלת עלות מינימאלית בין שני תתי עצים מינימאליים אשר איחודם כולל את כל הצמתים, תיתן עץ פורש מינימאלי.

הוכחה: יהיה S תת עץ אחד משני תתי העצים, ו V-S תת העץ השני. תהי $e=(v,w)$ קשת בעלת עלות מינימאלית אשר צידה האחד ב S וצידה השני ב V-S. לפי 4.17 קשת זו מוכלת בעץ פורש מינימאלי, ולכן אם נוסיפה אל שני תתי העצים אשר איחודם פורש נקבל עץ פורש מינימאלי כנדרש.

האלגוריתם:

1. חפש את הקשת המינימאלית ב `G המקשרת בין שני תתי העצים שהתקבלו מהסרת e מעץ T (לפי טענה 2)
2. הוסף את הקשת ל T וקבל `T הפורש את `G (טענה 1)

סיבוכיות:

מציאת קשת מינימאלית בין שני תתי העצים תיקח $O(|E|)$, ואילו הוספה שלה תיקח *O(1)*, ולכן הסיבוכיות יוצאת $O(|E|)$, כנדרש.

שאלה 2

יהי $G=(V,E), |V|>1$ גרף בעל משקלים אי שליליים

טענה 1:. יהי $x\in V$, אם נעשה דייקסטרה מצומת x, כל מסלול קצר ביותר יעבור בקשת אחת בלבד המקיימת $(x,u), u\in V$, ויעבור בה בדיוק פעם אחת.
הוכחה: נניח כי יש מסלול זול ביותר בין x ל y כלשהו, כך שבמסלול עוברים בשתי הקשתות, $(x,u),(x,v)$, מכאן שחייבת להיות במסלול קשת כלשהי $(t,x), t\in V$, ולכן, ללא הגבלת הכלליות נוכל למצוא מעגל, *x-u→u-t→t-x*.
לו נסיר את המעגל מהמסלול ונצא ישירות דרך קשת *(x,v)* המסלול יהיה זול יותר, בסתירה לכך שהמסלול שלנו הוא הזול ביותר. במקרה ש *u=v*, אז אותו הטיעון בדיוק יראה כי ניתן לוותר על המעברים המיותרים ולקבל מסלול זול יותר בסתירה. לסיום, אם נניח כי יש מסלול קצר ביותר שאינו עובר דרך אף קשת $(x,u), u\in V$, הרי שאין שום דרך לצאת מ *x* וזו סתירה לקיום מסלול כזה. ולסיכום הראנו כי עלינו לעבור בקשת אחת כזו פעם אחת בדיוק.

נאמר עתה כי בגרף G ישנה קשת שלילית אחת ויחידה, והיא $e=(x,u), w(e)<0$

טענה 2: הוספה של הערך |w(e)| לכל הקשתות $(x,v), v\in V$ והרצה של דייקסטרה על צומת x, תיתן מרחקים מינימאליים לכל צומת מ x הגדולים כולם בדיוק ב |w(e)| מהמשקלים האמיתיים.
הוכחה: ראשית, אם הוספנו את הערך |w(e)| לכל הקשתות $(x,u), u\in V$, הרי שקיבלנו חזרה גרף בעל משקלים אי שליליים עליו דייקסטרה מוגדר כראוי. לפי טענה 1, אנו מוכרחים לעבור דרך קשת שהוספנו לה |w(e)| פעם אחת ויחידה, ומהעובדה כי דייקסטרה עובד נכון, קיבלנו כי המשקלים יהיו גדולים בדיוק ב |w(e)| מהמשקל המקורי.

טענה 3: יהיו הנתונים הבאים:
1. P_{st} - המסלול הקצר ביותר מ s ל t העובר דרך קשת e=(x,y),
2. J_{sx} - המסלול הכי קצר מ s ל x
3. K_{xt} - המסלול הכי קצר מ x ל t

מתקיים $l(P_{st})=l(J_{sx})+l(K_{xt})$ (l היא פונקצית המרחק)
הוכחה: נניח בשלילה כי $l(J_{sx})+l(K_{xt})<l(P_{st})$, הרי שמצאנו מסלול קצר יותר - בסתירה ל-1.
לכן, נשאר להראות כי לא יתכן כי $l(J_{sx})+l(K_{xt})>l(P_{st})$ ואמנם, לו זה היה המצב, הרי שאם היינו לוקחים את מסלול P_{st} וחותכים אותו מכל אחד מקצותיו עד לצומת *x* היינו מקבלים את המסלולים L_{sx} ו L_{xt}, אשר היו מקיימים $l(P_{st})=l(L_{sx})+l(L_{xt})$ ומכאן שהיינו מקבלים כי $l(J_{sx})+l(K_{xt})>\ l(L_{sx})+l(L_{xt})$ ולכן לפחות אחד מהתנאים הבאים חייב להתקיים $J_{sx}>L_{sx}$ או $K_{xt}>L_{xt}$, וכל אחד מהם היה סותר את העובדה כי הם המסלולים הקצרים ביותר (2,3).

לכן נוכל לבנות את האלגוריתם הבא:

יהי $G=(V,E)$ גרף בעל קשת אחת $e=(x,y), w(e)<0$.

1. הרץ דייקסטרה על גרף G מ s עד אשר תגיע לצומת x. (עד צומת זו הקשתות אי שליליים)
2. אם דייקסטרה הצליח להגיע ל t, שמור את המרחק ב m
3. הוסף את הערך השלילי לפי טענה 2, והרץ דייקסטרה על גרף G מ x, והפחת את הערך השלילי והוסף את הערך שקיבלנו ב x משורה 1 לערך שהתקבל משורה 2 עבור צומת t
4. החזר את המינימום בין m לבין הערך ששורה 3 נתנה לצומת t

נכונות: אם המסלול הקצר ביותר כלל לא עבר דרך e, הרי ששורה 1 תמצא אותו ותשמור את הערך ב m.
אחרת, המסלול הקצר ביותר יעבור דרך e, ולכן שורה 1 תמצא את המרחק עד ל x, מטענה 2, נמצא את המרחק המדויק מ x ל t, ומטענה 3 הסכום שנקבל הוא המרחק אותו אנו מחפשים.

סיבוכיות: אנו מריצים דייקסטרה פעמיים, ולכן הסיבוכיות היא בסדר גודל של דייקסטרה, כלומר $O(|E|\ +\ |V|\log|V|)$.

שאלה 3

יהי $G = (V, E)$ גרף לא מכוון קשיר ללא משקלים.

נהפוך את גרף G לגרף מכוון `G, על ידי החלפת כל קשת (u, v) בזוג קשתות $(u, v), (v, u)$.

עתה, ניתן לכל קשת בגרף $e = (u, v)$ את המשקל $w(e) = rank(v)$

טענה 1: סכום של מסלול מ s ל t בגרף `G יהיה שווה לסכום דרגות של הקודקודים במסלול השקול בגרף G.
הוכחה: יהי $e_1, e_2, \dots, e_n | e_i = (a_i, b_i) \in E$ מסלול מ $\mathrm{a}_1 = s$ ל $b_n = t$ בגרף `G.
על פי ההגדרה של `G, קשת e_i מכוונת מ a_i ל b_i ומשקלה $w(e_i)$ הוא $rank(b_i)$, וכן צומת b_i זהה בשני הגרפים.
מכאן, מתקיים $\sum_{i=1}^{n} rank(b_i) = \sum_{i=1}^{n} w(e_i)$ וקיבלנו את הדרוש.

לפי טענה זו נבנה את האלגוריתם הבא:

1. הפוך את גרף G לגרף G` כמתואר לעיל
2. הרץ דייקסטרה על G`
3. המסלול שקיבלנו ב G` יהיה המסלול המינימאלי, ולפי טענה 1, הוא זהה למסלול שנקבל ב G עבור המסלול בעל סכום הדרגות המינימאלי.

סיבוכיות: להפוך את G ל G` ירוץ בזמן $\mathrm{O}(|E|)$, אולם הרצת דייקסטרה עולה יותר, ומכאן שהסיבוכיות זהה לסיבוכיות של דייקסטרה.

שאלה 4

א. יהי $G = (V, E)$ גרף קשיר ולא מכוון, ויהי $H = (V, F)$ גרף המתקבל מהאלגוריתם המתואר.
יהיו $a, b \in V$, נראה כי המסלול הקצר ביותר *a-b* ב *H* לא עולה על פי 3 מהמסלול הקצר ביותר ב *G*.

טענה 1: אם $e = (x, y) \notin F$, אז קיים מסלול x-y ב H, כך שמתקיים $d_{xy} \leq 3l_e$
הוכחה: טריוויאלית, האלגוריתם תוכנן באופן מפורש להכניס את קשת e ל F אם קיים מסלול x-y ב H כך שאורכו עולה על $3l_e$, ולכן בטוח יהיה מסלול הקטן מ $3l_e$.

טענה 2: קיים מסלול x-y ב H אשר אורכו לא עולה על פי שלושה מהמסלול הקצר ביותר ב G
הוכחה: נסמן $x = v_1, y = v_{k+1}$, ויהי $\mathrm{M} = (v_1, v_2, \dots, v_{k+1})$ המסלול הקצר ביותר $v_1 - v_{k+1}$ בגרף G.
ונסמן $\left(e_1 = (v_1, v_2), e_2 = (v_2, v_3), \dots, e_k = (v_k, v_{k+1})\right)$ את כל הקשתות המרכיבות את מסלול M.
במקרה הגרוע ביותר, אף אחת מהקשתות e_i אינן ב-F, ולכן נניח כי זה המקרה (כל מקרה אחר יהיה טוב יותר). משקל המסלול הקצר ביותר ב G הינו $\sum_{i=1}^{k} l_{e_i}$, ולכן, נפעיל את טענה 1 על כל מחובר בסכום, ונקבל כי לכל קשת e_i כזו, קיים מסלול ב H אשר ארכו אינו עולה על פי 3, כלומר

$$d_{xy} = d_{v_1 v_2} + \cdots + d_{v_k v_{k+1}} \leq 3l_{e_1} + \cdots + 3l_{e_k} = 3 \sum_{i=1}^{k} l_{e_i}$$

לכן, לכל זוג צמתים מ *V*, נוכל לפי טענה 2 למצוא מסלול ב-*H* אשר קטן מהמסלול הקצר ביותר ב *H* פי שלוש, ולכן לא יתכן כי יהיה מסלול קצר ביותר ארוך יותר ב *H*, שכן מצאנו מסלול קצר ממנו.

כנדרש.

- דרך נוספת להוכיח היא בעזרת טיעון החלפה, היינו מוצאים את הקשת הראשונה במסלול ב *H* אשר שונה מהמסלול ב *G* ואז לפי טענה 1 מובן שקיימת דרך חלופית בעלות שאינה עולה על $3l_e$..

ב. **טענה 1 :** קשת כלשהי תצטרף לגרף H בתנאי שהמסלול הקיים עד כה גדול מ-3 קשתות

הוכחה : בשלילה נניח כי האלגוריתם הוסיף ל H קשת $e_4=(u,v)$, כאשר כבר קיים מסלול u-v כך שמספר הקשתות שלו קטן שווה ל-3. נניח שמספר הקשתות שווה ל-3, (עבור 2 ו 1 ההוכחה אנלוגית להוכחה זו), נסמן ב (e_1,e_2,e_3) את המסלול u-v, וללא הגבלת הכלליות נאמר $l_{e_1}<l_{e_2}<l_{e_3}$.

מכך שהאלגוריתם הוסיף את קשת e_4 ל H, משתמע כי מתקיימות התכונות הבאות:

1. $l_{e_1}<l_{e_4}$, אחרת האלגוריתם היה מכניס את l_{e_4} קודם.
2. $3l_{e_4}<l_{e_1}+l_{e_2}+l_{e_3}$, אחרת, האלגוריתם לא היה מכניס את l_{e_4} כלל.

נתון כי משקל כל הקשתות שונה, ולכן, קיימים $0<a<b<c$ כך ש

$$3l_{e_4}=3\left(l_{e_1}+c\right)<l_{e_1}+l_{e_2}+l_{e_3}=l_{e_1}+\left(l_{e_1}+a\right)+\left(l_{e_1}+b\right)=3l_{e_1}+a+b$$

כלומר $3l_{e_1}+3c<3l_{e_1}+a+b$

ומכאן $3c<a+b$

אולם $a+b<b+b<c+b<c+c<3c$ ולכן קיבלנו סתירה!

טענה 2 : כל מעגל C ב H, מכיל יותר מ 4 קשתות.

הוכחה : לפי טענה 1, קשת אשר תיצור מעגל ב H חייבת להצטרף למסלול המקשר בין שני צמתי הקשת בעל 4 קשתות לפחות, ולכן המעגל שייווצר יהיה בעל יותר מ 4 קשתות.

טענה 3 : בגרף H בעל n צמתים הנוצר מהאלגוריתם המתואר, הדרגה המינימאלית קטנה מ $\sqrt{n}$

הוכחה : בשלילה נניח כי הדרגה <u>המינימאלית</u> גדולה מ $\sqrt{n}$, כלומר יש צומת $v\in H$ בעלת $\sqrt{n}$ קשתות לפחות אשר מחברות את v עם $\sqrt{n}$ צמתים שונים. מטענה 2 ברור כי אף אחד מ $\sqrt{n}$ הצמתים הללו לא מחוברים ביניהם, שכן אחרת היה נוצר מעגל בעל 3 קשתות. באותו האופן, לכל אחד מ $\sqrt{n}$ הצמתים הללו יש לפחות $\sqrt{n}-1$ קשתות (קשת אחת הולכת ל-v) עבור צמתים שונים, וגם כאן הצמתים לא יכולים להיות מחוברים בניהם, אחרת היינו מקבלים מעגל בעל 4 קשתות בסתירה לטענה 2.

נסכם את כמות הצמתים שמנינו עד כה: $1+\sqrt{n}+\sqrt{n}\left(\sqrt{n}-1\right)=1+\sqrt{n}+n-\sqrt{n}=n+1$

בסתירה לנתון כי ב H רק n צמתים. מש"ל.

טענה 4 : בגרף H בעל n צמתים הנוצר מהאלגוריתם המתואר, אין יותר מ $n^{1+\frac{1}{2}}$ קשתות.

הוכחה : מטענה 3, קיימת ב H צומת v בעלת $\sqrt{n}$ קשתות לכל היותר. ניצור גרף חדש, H_1, על ידי הסרת צומת v וכל הקשתות שלה ($\sqrt{n}$ קשתות). מכאן, H_1 הוא גרף בעל n-1 צמתים וללא מעגלים הקטנים מ 5 קשתות (כי יצרנו אותו על ידי הסרה של קשתות בלבד), ולכן מטענה 3 יש לו צומת u בעלת $\sqrt{n-1}$ קשתות לכל היותר, ולכן ניצור גרף H_2 על ידי הסרת הצומת u ולו n-2 צמתים וחסר מעגלים הקטנים מ 5 כמו קודם.

נמשיך כך עד שנגיע לגרף H_n אשר יהיה גרף ריק. בסך הכל, הסרנו $\sum_{i=1}^{n}\sqrt{i}$ קשתות לכל היותר.

ולכן, מתקיים: $f(n)\le\sum_{i=1}^{\mathrm{n}}\sqrt{i}<\sum_{i=1}^{n}\sqrt{n}=n\sqrt{n}=n^{1+\frac{1}{2}}$ ובכך חסמנו את מספר הקשתות ,כנדרש.

ולשאלה עצמה: מטענה 4 נוכל לומר:

$$\frac{f(n)}{n^2}<\frac{n^{1+\frac{1}{2}}}{n^2}=\frac{1}{\sqrt{n}}\xrightarrow[n\to\infty]{}0$$

ונקבל

כנדרש.

<u>שאלה 5</u>

יהי *T* עץ בינארי לחלוטין בעל *n* עלים, ובעל עומק *d*.

ניצור *d* קבוצות, כך ש $L_d = \{x | x \in T, depth[x] = d\}$

עתה, לכל עלה בכל קבוצה ניתן את הערך $\forall L_i, 1 \le i \le d, \forall x \in L_i, w(x) = \frac{1}{2^d}$:

טענה 1 : עבור כל *d>0*, ברמה התחתונה ביותר יש מספר זוגי של עלים
הוכחה : מכך ש *d>0*, ברור כי בעץ יש יותר מצומת אחת. נניח בשלילה כי קיימת קבוצה בה יש מספר אי זוגי של עלים ברמה התחתונה ביותר, מכאן שיש עלה אחד שאין לו אח, בסתירה לנתון כי *T* עץ בינארי לחלוטין.

תהי $F = \{w(f_i) | f_i \in \cup L_d\}$ סדרת השכיחויות הרצויה.

טענה 2 : סכום כל האיברים ב *F* שווה ל 1 בדיוק.
הוכחה : ניקח את כל זוגות העלים-אחים שנתקבלו מהרמה התחתונה ביותר, L_d. מטענה 1 ברור כי ישנם מספר זוגי של עלים, וכן סכום כל זוג אחים שווה לסכום של עלה ברמה מעל: $\frac{2}{2^d} = \frac{2}{2*2^{d-1}} = \frac{1}{2^{d-1}}$. עבור כל זוג כאלו, נוסיף עלה עם ערך $\frac{1}{2^{d-1}}$ לקבוצה מעל ונמחק את הקבוצה הזו. מכאן שהקבוצה מעל, L_{d-1}, הופכת להיות הנמוכה ביותר, ולכן נחזור שוב על התהליך עד אשר נגיע לקבוצה L_0 אשר מכילה רק צומת יחיד, ובפרט ערכו 1.

טענה 3 : סדרת השכיחויות *F* תיתן עץ הופמן *T*.
הוכחה : למעשה, נראה כי האלגוריתם של הופמן ייצר בדיוק את עץ *T*. בכל איטרציה האלגוריתם ישתמש בזוג השכיחויות הנמוך ביותר,והזוג הנבחר תמיד יגיע מהקבוצה הנמוכה ביותר, כלומר בזוג העלים הנמוכים ביותר בעץ *T* המקורי. בדומה להוכחת טענה 2, האלגוריתם יאחד את הזוג הנמוך ביותר ויוסיף את האיחוד אל הקבוצה מעל. הערך של האיחוד הנ"ל יהיה זהה בדיוק לשאר האותיות בקבוצה מעל, שוב בגלל אופי הבחירה של הערכים שלנו. בכל איחוד כזה האלגוריתם של הופמן יצור שורש חדש לזוג שנבחר, והשורש הנ"ל יהיה למעשה ה"אות" החדשה ברמה שמעל. באיטרציה האחרונה נקבל את שורש העץ, אשר לפי טענה 2 יהיה שווה ל-1 (שזה מה שאלגוריתם הופמן דורש), והוא יהיה בדיוק השורש של העץ *T* ממנו התחלנו.

כנדרש.