אלגוריתמים – פתרון ממ"ן 12

שאלה 1

תיאור האלגוריתם:

בכל שלב נבחר לתזמן המשימה שעבורה הערך $\frac{w_i}{l_i}$ הוא הגדול ביותר, כלומר התזמון יהיה המשימות ממוינות לפי $\frac{w_i}{l_i}$, מהגדול לקטן. הסיבה לכך שהיא שנרצה לתזמן קודם את המשימה עם המשקל הגדול ביותר, אך נרצה שאורכה יהיה כמה שיותר קצר. היחס הזה הוא בעצם המשקל ביחס לזמן, והמשימה שמשקלה הכי גדול בזמן הקצר ביותר היא המשימה הטובה ביותר לבחירה בכל רגע. נכנה את יחס זה בשם ערך המשימה.

הוכחת נכונות:

יהי O פתרון אופטימלי לפתרון הבעיה. נרצה להוכיח שסידור המשימות של O מקיים $\frac{w_i}{l_i} \leq \frac{w_j}{l_j}$, לכל שני משימות עוקבות i, j כך ש- j מתוזמנת לפני i. נניח בשלילה שהסידור שמפיק O לא ממוין באופן זה. כלומר קיימות משימות עוקבות $1 \leq a,b \leq n$ כך שהמשימה a מתוזמנת לפני המשימה b, וש- $\frac{w_a}{l_a} < \frac{w_b}{l_b}$. נניח שהמשימה a מתוזמנת כך שתתחיל בזמן s. משימה זו מסתיימת בזמן $s+l_a$. המשימה b מתחילה בזמן $s+l_a$ ומסתיימת בזמן $s+l_a+l_b$.

נתבונן באלגוריתם O' שבו מוחלפות המשימות a ו- b. באלגוריתם זה b מתחילה בזמן s, מסתיימת בזמן $s+l_b$. המשימה a מתחילה בזמן $s+l_b$ ומסתיימת בזמן $s+l_a+l_b$. לא שינינו את זמני ההתחלה והסיום של צמד המשימות a ו- b ולכן לא השפענו על המשימות האחרות. O אופטימלי, ולכן התזמון שהוא מפיק חייב להיות טוב לפחות כמו התזמון שמפיק O', כלומר:

$\sum_{i=1}^{n} w_i t_i \leq \sum_{j=1}^{n} w_j t'_j$, כאשר t'_i הוא זמן הסיום בתזמון ש- O' מייצג, ו- t_i הוא זמן הסיום בתזמון ש-

O מייצג. מלבד a ו- b, כל המשימות האחרות מזומנות באותם זמנים ובאותו סדר, ולכן עבור משימות אלו $w_i t_i = w_j t'_j$ עבור $i, j \neq a,b$. נמשיך:

$$\begin{aligned}
&\sum_{\substack{i=1 \\ i\neq a,b}}^{n} w_i t_i + w_a t_a + w_b t_b < \sum_{\substack{j=1 \\ j\neq a,b}}^{n} w_j t'_j + w_a t'_a + w_b t'_b \Rightarrow \\
&w_a t_a + w_b t_b < w_a t'_a + w_b t'_b \Rightarrow \\
&w_a\left(s+l_a\right) + w_b\left(s+l_a+l_b\right) < w_b\left(s+l_b\right) + w_a\left(s+l_a+l_b\right) \Rightarrow \\
&w_a s + w_a l_a + w_b s + w_b l_a + w_b l_b < w_b s + w_b l_b + w_a s + w_a l_a + w_a l_b \Rightarrow \\
&w_b l_a < w_a l_b \\
&\frac{w_b}{l_b} < \frac{w_a}{l_a}
\end{aligned}$$

וזוהי סתירה להנחה.

מכאן נובע שבסידור שמפיק O כל שתי משימות עוקבות מסודרות לפי המשקל ביחס לזמן, ולכן התזמון של O מכיל משימות מסודרות לפי $\frac{w_i}{l_i}$, מהגדול לקטן. התזמון של O יכול לפיכך, להיבדל מזה של האלגוריתם שלנו רק בסדר של משימות השוות בערכן, אבל קל לאמת בדרך שלמעשה שקולה למעבר הארוך לעיל שעבור משימות ששוות בערכן a,b, המתוזמנות בסדר הפוך בשני האלגוריתמים:

$$\begin{aligned}
&\frac{w_a}{l_a} = \frac{w_b}{l_b} \Rightarrow w_a l_b = w_b l_a \Rightarrow w_a\left(s+l_a\right) + w_b\left(s+l_a+l_b\right) = w_b\left(s+l_b\right) + w_a\left(s+l_a+l_b\right) \\
&\Rightarrow w_a t_a + w_b t_b = w_a t'_a + w_b t'_b
\end{aligned}$$

כאשר t' מייצג את תזמוני האלגוריתם שלנו. הווה אומר שהסדר של שני משימות השוות בערכן לא משפיע על הסכום לו רוצים למצוא מינימום, ומכיוון שזה ההבדל היחיד בין הסידור שיפיק האלגוריתם שלנו לבין הסידור של אלגוריתם אופטימלי, נובע שהאלגוריתם שלנו מחזיר תזמון בעל סכום השווה לזה של האלגוריתם האופטימלי, כלומר את הסכום המינימלי.

ניתוח סיבוכיות:

נצמיד לכל משימה את ערכה בזמן $O(n)$. נמיין את המשימות לפי ערכן בעזרת מיון השוואות אופטימלי, בזמן $O(n\log n)$. נחזיר את מספרי המשימות לפי סדר המערך הממוין.

בסך הכל זמן הריצה הוא $O(n\log n)$.

<u>שאלה 2</u>

תיאור האלגוריתם:

ניצור גרף חדש על ידי כך שנשנה את משקלי הקשתות שבקבוצה F. נניח ש- $w:E\to\mathbb{R}_{non-negative}$ היא פונקציית המשקל המקורית על הקשתות ב- G. ניצור פונקציית משקל חדשה $c:E\to\mathbb{R}_{non-negative}$ לכל $e\in F$ נקבע $c(e)=0$, ולכל $e\notin F$ נקבע $c(e)=w(e)+1$.

למעשה איפסנו את משקלי קשתות היער הנתון, כך שהוא יימצא בעץ הפורש שנמצא תוך שימוש בפונקצית המשקל החדשה.

לאחר מכן נריץ את אחד האלגוריתמים למציאת עץ פורש מינימלי על הגרף G עם פונקציית המשקל c ונחזיר את העץ המתקבל.

הערה: דרך אחרת לפתור את הבעיה היא לבצע את האלגוריתם של קרוקסאל, כאשר מאתחלים את היער שנבנה על ידי האלגוריתם ב- F. למעשה זה מקרה פרטי של הדרך לעיל, שכן אם משקלי היער הנתון הם אפס, האלגוריתם יצרף אותם ליער בהתחלה.

הוכחת נכונות:

יש להוכיח את הטענות הבאות:

(1) העץ המתקבל הוא עץ פורש המכיל את כל קשתות היער F.

(2) העץ המתקבל הוא המינימלי מבין כל העצים הפורשים אשר מכילים את קשתות היער F.

<u>הוכחות:</u>

(1) מכיוון שאנו מריצים אלגוריתם למציאת עץ פורש מינימלי על הגרף המקורי, ורק שינינו את העלות של הקשתות, העץ שיוחזר יהיה גם עץ פורש עבור הגרף G.

נניח בשלילה שהעץ המתקבל אינו מכיל את כל קשתות היער F. כלומר קיימת קשת $e\in F$ כך ש- e לא נמצאת ב- T. F הוא יער, ולכן חסר מעגלים. הוספת הקשת e ל- T תיצור מעגל, אך לא ייתכן שכל הקשתות על המעגל תהיינה קשתות של F, משום שאז ב- G' יש מעגל בסתירה לכך שהוא יער. מכאן קיימת על המעגל קשת f שמשקלה

$c(f)$ הוא חיובי; ההחלפה $T'=T-\{f\}\cup\{e\}$ תתן עץ פורש שמשקלו קטן יותר מזה

של T (ראה נספח), עבור פונקציית המשקל c, ולכן קיבלנו סתירה.

(2) יהי T עץ פורש של G, שכולל את כל קשתות F. נתבונן בהשפעת החלפת פונקציית המשקל על משקלו של T. לפי הגדרת c, מתאפסים (ולכן יורדים) כל המשקלים של הקשתות שביער F, ונוסף 1 עבור כל קשת שאינה ב- F (כלומר נוסף סך כל מספר הקשתות שב- T אבל לא ב- F). לפיכך: $c(T)=w(T)-w(F)+\left(|E(T)|-|F|\right)$. עבור כל עץ מתקיים $|E(T)|=|V|-1$, ובסך הכל נקבל:

$c(T)=w(T)-w(F)+\left(|V|-|F|-1\right)$.

כעת נניח ש- T הוא העץ שהאלגוריתם החזיר. נניח בשלילה שקיים עץ T' כך ש-T' כולל את כל קשתות F וגם $w(T')<w(T)$. T עץ פורש מינימלי ביחס ל- c, כלומר $c(T)\le c(T')$. מהאמור לעיל נובע ש-

$w(T)-w(F)+\left(|V|-|F|-1\right)\le w(T')-w(F)+\left(|V|-|F|-1\right)$ ולכן $w(T)\le w(T')$ - הגענו לסתירה.

ניתוח סיבוכיות:

בשלב הראשון נצטרך לעדכן את משקלי כל הקשתות, ולכן צעד זה ייקח $O(|V|)$.

מציאת העץ הפורש המינימלי עם פונקציית המשקל החדשה יכולה להעשות בזמן $O(|E|\log|V|)$, למשל על ידי האלגוריתם של פרים או האלגוריתם של קרוסקאל.

בסך הכל מדובר בזמן ריצה של: $O(|E|\log|V|)+O(|V|)=O(|E|\log|V|+|V|)$. הגרף קשיר ולכן $|E|\ge|V|-1$, כלומר $|V|=O(|E|)$, ונקבל בסך הכל: $O(|E|\log|V|+|E|)=O(|E|\log|V|)$.

שאלה 3

יהי $T=\left(V,E'\right)$ עץ פורש מינימלי של G. נניח בשלילה שקיים ב-G צומת u עם דרגה גדולה או שווה ל-7. לכן קיימות לפחות 7 קשתות שיוצאות מ-u. קשתות אלו מחלקות את המעגל שמרכזו ב-u ללפחות 7 גזרות. סכום הזוויות בין הישרים שווה ל-360 מעלות, מכיוון שהן מקיפות מעגל שלם. לפחות אחת מן הזוויות קטנה מ-60 מעלות: אם כל הזוויות בנות לפחות 60 מעלות, אז סכומן לפחות $7\cdot 60=420>360$, בסתירה לכך שסכום הזוויות המרכזיות במעגל הוא 360 מעלות.

נניח שהזווית הקטנה מ-60 המעלות נמצאת במשולש המורכב מהצמתים x, y וכמובן u, כאשר $\angle xuy<60^\circ$. לפי משפט בגיאומטריה, מול הזווית הקטנה ביותר במשולש, נמצאת הצלע הקטנה ביותר במשולש (נובע למשל ממשפט הסינוסים: $\frac{a}{\sin\alpha}=\frac{b}{\sin\beta}=\frac{c}{\sin\gamma}$, ולכן אם $\alpha<\beta,\gamma$ אז בהכרח $a<b,c$). הווה אומר, ש-$w(x,y)<w(u,x),w(u,y)$ (1).

יהי $T'=\left(V,E'-\{u,x\}\cup\{x,y\}\right)$. נוכיח ש-$T'$ עץ פורש מינימלי של G:

לא שינינו את מספר הקשתות, לכן עדיין מתקיים $\left|E_{T'}\right|=\left|V\right|-1$. T' קשיר: בעץ המקורי יש מסלול מכל צומת לכל צומת. יהיו a,b צמתים ב-T'. נתבונן במסלול בינהם בגרף T. אם המסלול מכיל את הקשת $\{u,x\}$, נחליף אותה בקשתות $\{u,y\},\{y,x\}$. המסלול החדש נמצא בגרף T', ולכן הגרף קשיר, כלומר T' עץ (שכן $\left|E_{T'}\right|=\left|V\right|-1$). כמובן ש-$T'$ עץ פורש עבור הגרף G, כי קבוצות הצמתים שלהם שווים.

מתקיים: $w(T')=w(T)+\overbrace{w(x,y)-w(u,x)}^{<0\quad(1)}<w(T)$. מכאן נובע שהמשקל של T' קטן מזה של T, בסתירה לכך ש-T עץ פורש מינימלי.

שאלה 4

ראשית כל, נבחין בכך שאם משקלי הקשתות בגרף לא שונים אנו נתקלים בבעיה בפתרון השאלה. אם הכוונה בעץ הפורש המינימלי השני הטוב ביותר היא לעץ פורש שמשקלו קטן ממש מזה של T, אז ייתכן שאין כזה, למשל במצב בו כל משקלי הקשתות שווים ולכן כל עץ הוא עפ״מ. לכן נניח שהכוונה היא לעץ שמשקלו יכול להיות שווה לזה של T.

יהי T עץ פורש מינימלי של G. הגרף G מקיים $|E| \geq |V|$, ולכן אפשר לומר שקיימים עוד עצים על G ששונים מ- T. אם קיימים בגרף שני עפ״מים שונים, אז הם מכילים קשתות שמשקליהן שווה על מעגל, שנמצאות על פי הוכחת (**) בעפ״מים. אם נחליף ב- T את אחת הקשתות בשנייה, נקבל עפ״מ שמשקלו שווה לזה של T, ולכן נוכל לכנות אותו ״שני״.

אם כן, נניח שבגרף אין שני עפ״מים שונים.

יהי T_2 עץ כך ש- T_2 נבדל מ- T ב-2 או יותר קשתות, כלומר החלפה ב- T_2 של הקשתות שב- $T_2 - T$ בקשתות אחרות שנמצאות ב- $T - T_2$ תחזיר אותנו ל- T. נניח בשלילה נניח ש- T_2 העץ הפורש המינימלי השני הכי טוב של G. קיימת ב- $T_2 - T$ קשת $\{x, y\}$. נשים לב שאם נצרף את $\{x, y\}$ לעץ T, יווצר מעגל (כי הגרף יישאר קשיר, אך מספר הקשתות יהיה גדול או שווה ממספר הצמתים).

מעגל זה יכיל את הקשת $\{x, y\}$. כמו כן, לא כל קשתות המעגל נמצאות ב- T_2, כי הוא עץ, ולכן קיימת על המעגל, קשת $\{u, v\}$ מ- T, כך ש- $\{u, v\} \notin T_2$. אם $w(\{u, v\}) = w(\{x, y\})$, אז קיבלנו שני קשתות שוות על אותו מעגל, ולכן אם נחליף את אחת הקשתות בשנייה נקבל שיש לעץ שני עפ״מים.

אם כן, נסיק ש- $w(\{u, v\}) \neq w(\{x, y\})$. נפריד למקרים:

- אם $w(\{u, v\}) > w(\{x, y\})$, אז נתבונן בגרף T' בו מחליפים את הקשת $\{u, v\}$ בקשת $\{x, y\}$. לפי הנספח, T' הוא עץ. אך משקלו של T' קטן מזה של T, כי החלפנו בו קשת כבדה אחת, בקשת קלה יותר.

- אחרת, $w(\{u, v\}) < w(\{x, y\})$. נתבונן בגרף T', שמתקבל בהחלפת הקשת $\{x, y\}$ בקשת $\{u, v\}$ בגרף T_2. לפי הנספח, T' עץ. מתקיים $w(T') = w(T_2) + \overbrace{w(\{u, v\}) - w(\{x, y\})}^{<0} < w(T_2)$.

T' מכיל קשת $\{s',t'\}$ שלא נמצאת ב- T (כי T_2 מכיל לפחות שני קשתות שאינן ב- T, ואת T' בנינו כך שיהיו בו כל הקשתות של T_2 מלבד קשת אחת). לפיכך $T' \neq T$. אבל ל- G יש רק עץ פורש מינימלי יחיד אפשרי, והוא T. לכן T' הוא עץ פורש שמשקלו גדול מזה של העץ הפורש המינימלי, אך משקלו קטן מזה של T_2, וזה עומד בסתירה לכך ש- T_2 הוא העץ הפורש המינימלי השני.

הגענו למסקנה שכל עץ שנבדל מהעץ הפורש המינימלי T **בלפחות** שתיים מן הקשתות הוא לא העץ הפורש המינימלי השני הטוב ביותר. מכיוון שהעפ״מ השני הטוב ביותר בהכרח שונה מ- T, נקבל שהוא נבדל ממנו בדיוק בקשת אחת. הווה אומר, שקיימת בעץ T קשת $\{u,v\}$, וקשת נוספת $\{x,y\}$ שאינה בעץ T, כך ש- $T-\{u,v\}\cup\{x,y\}$ הוא העץ הפורש המינימלי השני הטוב ביותר.

(כפי שאחד הסטודנטים העיר, השאלה טבעית יותר כאשר מניחים שכל משקלי הקשתות שונים, ואז ניתן להוכיח שאין שני עפ״מים על ידי אותה ההוכחה של (**)).

(**) <u>טענה:</u> אם קיימים שני עפ"מים, אז בגרף המקורי יש שני קשתות בעלות משקל שווה שנמצאות על אותו מעגל, ואחת מהן נמצאת בלפחות עפ"מ אחד.

<u>הוכחה:</u>

נניח שבגרף המקורי אין שני קשתות בעלות משקל שווה שנמצאות על אותו מעגל ונמצאות בלפחות עפ"מ אחד, ונוכיח שאין שני עפ"מים. נניח ש- S ו- S' הם שני עצים פורשים מינימליים שונים של G. לכן קיימת ב- S קשת $\{s,t\}$ שאינה נמצאת ב- S'. S ו- S' עצים ולכן מספר הקשתות בהם שווה. לכן גם ב- S' קיימת קשת $\{s',t'\}$ שלא נמצאת ב- S. אם נשים את הקשת $\{s',t'\}$ ב- S, יווצר מעגל. מלבד קשת אחת, כל הקשתות במעגל נמצאות באחד העפ"מים, ולכן לפי ההנחה משקלי הקשתות $\{s,t\}$ ו- $\{s',t'\}$ שונים. לכן אחד מבין המשקלים של $\{s,t\}$ ושל $\{s',t'\}$ קטן יותר מהשני. נניח בלי הגבלת הכלליות שמשקל הקשת $\{s',t'\}$ נמוך מזה של $\{s,t\}$. נתבונן בגרף $R = S - \{s,t\} \cup \{s',t'\}$. באופן דומה לטיעוני ההחלפה שבהוכחה לעיל, הגרף החדש שנוצר הוא עץ (פורש). אבל $w(R) = w(S) - w(s,t) + w(s',t') < w(S)$ (משום ש- $w(\{s',t'\}) < w(\{s,t\})$) ולכן R הוא עץ פורש שמשקלו נמוך יותר משל S, בסתירה לכך שהוא עץ פורש מינימלי.

שאלה 5

נבצע רדוקציה לבעיית קוד הופמן: כל רשימה תהיה אות בא״ב. השכיחות של כל אות תהיה אורך הרשימה המתאימה לה. מיזוג של שתי רשימות נותן רשימה שאורכה הוא סכום אורכי הרשימות, ולכן, כמו בבעיית קוד הופמן, השכיחות של הרשימה הממוזגת הוא סכום שכיחויות הרשימות. אחרי תרגום הקלט לקלט לבעיית קוד הופמן נפעיל את האלגוריתם של הופמן ונקבל עץ עם עלות מינימלית, כאשר עלות העץ סוכמת עבור כל אותיות הא״ב את מכפלת שכיחות האות בעומק האות בעץ. לכן, עלות העץ למעשה סוכמת עבור כל הרשימות את מכפלת אורכן במספר המיזוגגם שהן עוברות, כלומר, את עלות המיזוג. לכן, מנכונות האלגוריתם של הופמן, התקבל עץ המייצג מיזוגים שעלותם מינימלית. הסיבוכיות היא כמובן כמו סיבוכיות האלגוריתם של הופמן, כלומר: $O(klogk)$. נשים לב, שאם לא נתון אורכה של כל רשימה, יש קודם לחשב זאת, אך זה ניתן לביצוע בזמן ליניארי באורך הקלט, כי הקלט כולל את הרשימות עצמן.

(*) **<u>נספח:</u>**

<u>טענה:</u>

נניח ש- G גרף, ו- T עץ פורש של G. נניח גם, שקיימות ב- G קשתות $e \in T$, $e' \notin T$, ושקיים ב- G מעגל C המכיל את e' ו- e. העץ T' בו מחליפים את e ב- e' הוא עץ פורש של G.

<u>הוכחה:</u>

נוכיח ש- T' קשיר: יהיו u,v צמתים ב- G. נתבונן במסלול $u-v$ בעץ T. אם המסלול עובר דרך הקשת e, אז במקום לעבור דרכה, ניתן לעקוף אותה דרך המעגל C: אם נניח בה"כ ש- $e=\{x,y\}$, ונגיע במסלול ל- x, אז x נמצא במעגל ולכן נעבור מ- x על המעגל, שקיים בעץ החדש, משום ששמנו בו את e', עד ל- y, ומשם נמשיך במסלול.

נוכיח ש- T' חסר מעגלים: המעגל היחיד ב- $T \cup \{e'\}$ הוא C, שכן לפני ההוספה ב- T אין מעגלים.

ממחיקת הצומת e נסיק שב- T' אין כלל מעגלים.

לפיכך, T' הוא עץ, ומכיוון שלא שינינו את קבוצת הצמתים, הוא גם עץ פורש.