פתרון בוחן בתכנון אלגוריתמים, 11.05.07

שאלה 1

סעיף א

הגדרת האיבר הכללי (בעברית)

$M[l,i]$ -

הרווח המקסימאלי שניתן להפיק מסרט באורך l ומההזמנה $Order_i = \langle (n_1,v_1),(n_2,v_2),...,(n_i,v_i) \rangle$

נוסחת רקורסיה ותנאי בסיס

תנאי בסיס: אם אין שום פריט בהזמנה, הרווח הוא 0 ולכן: $M[l,0]=0 \quad \forall\, 0 \le l \le L$

אם אורך הסרט הוא 0, לא ניתן לגזור ממנו כלום והרווח הוא 0 ולכן: $M[0,i]=0 \quad \forall\, 0 \le i \le m$

נוסחת נסיגה $M[l,i]=\max\{M[l-n_i,i-1]+v_i,\ \ M[l,i-1]\}$.

כלומר, בחירת מקסימום מבין שני ערכים כאשר הראשון מייצג מקרה בו מספקים סרט במידה n_i והשני מייצג מקרה בו לא מספקים סרט במידה n_i.

נימוק נכונות נוסחת הרקורסיה

צריך להראות שקיים תת-מבנה אופטימאלי כמו בנוסחת הנסיגה. יהי U פתרון אופטימאלי לבעיה עם אורך סרט l והזמנת קטעים $1,...,i$. נסמן ב- r_i את קטע הסרט ה-i-י **בהזמנה**. יש בדיוק שני מקרים.

אם $r_i \notin U$, הרי ש- U אופטימאלי גם עבור הבעיה עם סרט באורך l והזמנת קטעים $1,...,i-1$. אחרת, יהי $U'=U\setminus\{r_i\}$ תת-פתרון של U. נסמן ב- $len(U)$ את סך האורכים של קטעים ב- U. לפי הגדרת פתרון חוקי, הרי ש- $len(U')\le l-n_i$. נניח בשלילה ש- U' אינו אופטימאלי עבור הבעיה עם סרט באורך $l-n_i$ והזמנת קטעים $1,...,i-1$, כלומר קיים פתרון $U"$ טוב יותר. נתבונן בפתרון U^* (עבור הבעיה עם סרט באורך l והזמנת קטעים $1,...,i$) שמתקבל מ- $U"$ ע"י הוספה של r_i. קל לראות ש- U^* חוקי: $len(U")\le l-n_i$ ואורכו של r_i הוא בדיוק n_i, לכן $len(U^*)\le l$. כמו כן, מכיוון ש $U"$ מתאים להזמנת קטעים $1,...,i-1$, הרי שאינו מכיל את r_i, בסה"כ כל קטע מופיע ב- U^* לכל היותר פעם אחת. אבל $\text{cost}(U^*)=\text{cost}(U")+v_i>\text{cost}(U')+v_i=\text{cost}(U)$, בסתירה לאופטימאליות של U.

הערה- אין צורך להוכיח בבחינה ברמת פירוט כזו.

<u>אלגוריתם איטראטיבי</u>

Iter_Thread & Needle(L, order)

```
1   for l = 0 to L
2       M[l,0] = 0
3   for i = 0 to m
4       M[0,i] = 0
5   for l = 1 to L
6       for i = 1 to m
7           M[l,i] = M[l,i-1]
8           if l ≥ n_i and M[l-n_i, i-1] + v_i > M[l,i]
9               M[l,i] ← M[l-n_i, i-1] + v_i
```

הערה- אסור לחתוך מהסרט יותר מאשר את אורכו. לכן יש להתנות את ההתייחסות לערך $M[l-n_i,i-1]$ בכך ש- $l \ge n_i$.

<u>ניתוח זמן ריצה</u>

אתחול- שורות 1-2: $O(L)$ שורות 3-4: $O(m)$.

לולאה ראשית (כפולה)- $O(Lm)$, כי גוף הלולאה דורש O(1) בלבד.

סה"כ- $O(L)+O(m)+O(Lm)=O(Lm)$

<u>סעיף ב</u>

<u>הגדרת האיבר הכללי (בעברית)</u>

$M[l,i]$ -

הרווח המקסימאלי שניתן להפיק מסרט באורך l ומההזמנה $Order_i = \langle (n_1,v_1),(n_2,v_2),...,(n_i,v_i) \rangle$

<u>נוסחת רקורסיה ותנאיי בסיס</u>

תנאי בסיס כמו בסעיף א'.

<u>נוסחת נסיגה</u>: $M[l,i] = \max\{M[l-n_i,i]+v_i, \ \ M[l,i-1]\}$.

כלומר, בחירת מקסימום מבין שני ערכים כאשר הראשון מייצג מקרה בו מספקים עוד סרט במידה n_i והשני מייצג מקרה בו לא מספקים עוד סרט במידה n_i.

<u>נימוק נכונות נוסחת הרקורסיה</u>

צריך להראות שקיים תת-מבנה אופטימאלי כמו בנוסחת הנסיגה: יהי U פתרון אופטימאלי לבעיה עם אורך סרט l והזמנת קטעים $1,...,i$. נסמן ב- r_i את קטע הסרט ה-i-י **בהזמנה**. <u>יש בדיוק שני מקרים</u>.

אם $r_i \notin U$, הרי ש- U אופטימאלי גם עבור הבעיה עם סרט באורך l והזמנת קטעים $1,...,i-1$.

אחרת, נסמן ב- $len(U)$ את סך האורכים של קטעים ב- U, יהי U' תת-פתרון שמתקבל מ- U ע"י הסרה של קטע r_i אחד. לפי הגדרת פתרון חוקי, הרי ש- $len(U') \le l-n_i$. נניח בשלילה ש- U' אינו

אופטימאלי עבור הבעיה עם סרט באורך $l-n_i$ והזמנת קטעים $\boldsymbol{1,...,i}$, כלומר קיים פתרון U'' טוב יותר. נתבונן בפתרון U^* (עבור הבעיה עם סרט באורך l והזמנת קטעים $1,...,i$) שמתקבל מ- U'' ע"י הוספה של r_i. קל לראות ש- U^* חוקי: $len(U'') \le l-n_i$ ואורכו של r_i הוא בדיוק n_i, לכן $len(U^*) \le l$. (ומותר ל- U^* להכיל כמה מופעים של r_i).

אבל $\mathrm{cost}(U^*) = \mathrm{cost}(U'') + v_i > \mathrm{cost}(U') + v_i = \mathrm{cost}(U)$, בסתירה לאופטימאליות של U.

הערה- אין צורך להוכיח בבחינה ברמת פירוט כזו.

שאלה 2

<u>סעיף א</u>

<u>טענה ראשית</u>

עבור גרף G תהי C_e(G) קבוצת כל העצים הפורשים של G המכילים את הצלע e .
תהי A קבוצת הצלעות המוחזרת מן האלגוריתם. אזי תת-הגרף של G המושרה מ A, יהי זה T, שייך לקבוצה C_e(G) , ובעל משקל מינימלי מבין כל העצים בקבוצה C_e(G) .

<u>טענה נשמרת</u>

עבור גרף G נגדיר MST_e(G) להיות תת-קבוצה של C_e(G) כך שכל עץ ב MST_e(G) הינו בעל משקל מינימלי מבין העצים ב C_e(G).
תהי R(i) קבוצת הצלעות שלא נבחרו על ידי האלגוריתם לתוך A עד תחילת האיטראציה ה – i .
אזי בתחילת האיטראציה ה-i קיים עץ T בקבוצה MST_e(G), כך ש E(T) מכיל את A , ואינו מכיל אף צלע מ R(i) .

<u>טענת עזר</u>

יהיו :
- G=(V,E) גרף, ו- e צלע
- w:E→R פונקצית משקל על הצלעות,
- $T=(V,E_T)$ עץ פורש מינימאלי במשקלו מבין העצים הפורשים המכילים את הצלע e
- $e \in A \subseteq T$ קבוצת צלעות
- (S,V\S) חתך בגרף המכבד את A,
- $e' \in E$ צלע בעלת משקל מינימאלי מבין הצלעות החוצות את החתך הנתון

אזי :
קיים עץ פורש T_2 מינימאלי במשקלו מבין העצים הפורשים המכילים את e, כך ש $A \cup \{e'\} \subseteq T_2$.

<u>הוכחת הטענה הראשית</u>

בתום ריצת הלולאה נסקרו כל הצלעות בגרף ולכן R =E\A. על פי הטענה הנשמרת, קיים $T=(V,E_T)$ עץ בקבוצה MST_e(G) . אשר מכיל את צלעות A ואינו מכיל אף צלע מבין שאר הצלעות שנסקרו, כלומר מתוך R . לפיכך, מתקיים שוויון ממש (ולא רק הכלה) : $E_T=A$.
מ.ש.ל.

<u>הוכחת הטענה הנשמרת</u>

נתון כי בתום האיטרציה הקודמת, קיים T עץ פורש לגרף G המקיים את תנאי השאלה ומכיל את כל הצלעות שהוכנסו לקבוצה A, ואינו מכיל את הצלעות שנבחנו ולא הוכנסו לקבוצה A. באיטרציה הנוכחית נסקרת הצלע e'=(u,v) , וייתכנו שני מקרים:

מקרה א : $FindSet(u) = FindSet(v)$

אזי u ו-v נמצאים באותה קבוצה. קודקודים נכנסים לאותה קבוצה רק כאשר מוכנסת גם צלע לקבוצה A המחברת בין שתי הקבוצות. כלומר, כל קבוצה שנבנית במהלך האלגוריתם הינה רכיב קשירות על פי הצלעות ב-A (כפי שנלמד בכיתה). אי לכך, קיים מסלול בין u לבין v באמצעות צלעות A. אם נוסיף ל-A גם את הצלע e' , הרי שיהיו שני מסלולים שונים בין u לבין v , וברור שלא קיים עץ פורש המכיל את כל הצלעות הללו. בפרט T , אשר מכיל את צלעות A ואינו מכיל את הצלעות שנסקרו ולא הוכנסו ל-A, גם אינו מכיל את הצלע e', ולפיכך T מקיים את הטענה הנשמרת גם בתחילת הלולאה הבאה.

מקרה ב: $FindSet(u) \neq FindSet(v)$

אזי u ו- v אינם באותו רכיב קשירות על פי צלעות A. נגדיר חתך בגרף (S, V\S) כך ש- S יהיה רכיב הקשירות של u. אזי החתך מכבד את A (אחרת, היה מסלול בין u לבין קודקוד שאינו ברכיב הקשירות שלו באמצעות צלעות A. בסתירה). מאותה הסיבה, הצלע e'=(u,v) חוצה את החתך הזה (u ו-v אינם באותו רכיב הקשירות. רכיב הקשירות של u הינו צד אחד של החתך). הצלע e'=(u,v) הינה בעלת משקל מינימאלי מבין הצלעות החוצות את החתך. (אחרת, הייתה מופיעה צלע אחרת החוצה את החתך לפני e' בסדר המיון, והצלע הייתה מוכנסת על ידי האלגוריתם לקבוצה A. אז החתך לא היה מכבד את A). על פי טענת העזר קיים T* עץ פורש המקיים את תנאי השאלה ומכיל את $A \bigcup \{e'\}$.נשים לב, כי לא ייתכן והוא מכיל צלעות שנסקרו ולא הוכנסו לקבוצה A, משום שאז היו בו מעגלים, כפי שפורט במקרה א'. אי לכך, T* מקיים את האינווריאנטה גם בתחילת האיטרציה הבאה.

<u>הוכחת טענת העזר</u>

נציין MST_e של G להיות עץ פורש של G, מינימאלי במשקלו אשר מכיל את e.

אם T מכיל את e' , הרי ש- $A \cup \{e'\} \subseteq T$. נבחר $T_2 = T$, וסיימנו.

אם לא: יהי $T_1 = (V, E_T^1)$ גרף שנוצר ע"י הוספת e' ל- T, ז"א $E_T^1 = E_T \cup \{e'\}$. אם כן, ב-T_1 קיימים כעת שני מסלולים בין קודקודי הצלע e' (אחרת סתירה לכך ש T הוא עץ פורש). אי לכך, קיים מעגל ב-T_1 שמכיל את e' . כיוון שe' חוצה את החתך $(S, V \setminus S)$, אזי קיימת, במעגל זה, צלע e'' נוספת שגם חוצה את החתך (זאת כי ב.ה.כ נניח כי $u \in S, v \in V \setminus S$ ולכן מעגל שמתחיל ב- u, ממשיך ל- v חייב לחצות את החתך פעם נוספת כדי לחזור ל- u).

יהי $T_2 = (V, E_T^2)$ עץ שמתקבל מ- T ע"י הוספת e' והורדה ש ל e'' , ז"א.

$E_T^2 = E_T^1 \setminus \{e''\} = E_T \setminus \{e''\} \cup \{e'\}$

נראה כי:

1. T_2 הוא עץ פורש של G. **2.** T_2 מכיל את e ואת A. **3.** T_2 מינימאלי במשקלו.

1. מכיון שהצלע e' אינה בקבוצת הצלעות E_T^2, הרי שמבניית E_T^2 נובע: $\left|E_T^2\right| = \left|E_T\right| - 1 + 1 = \left|E_T\right| = n - 1$. בנוסף, נשים לב כי לאחר הוספת הצלע e'=(u,v) הגרף T_1 היה קשיר (כי T היה קשיר). טענו כי e'' נמצאת על מעגל ב- T_1 ולכן גם לאחר הסרתה, הגרף T_2 קשיר. אי לכך, T_2 עץ פורש.

2. נשים לב כי הצלע e'' שהסרנו לא שייכת ל-A, (ובהכרח $e'' \neq e$). זאת מכיון שצלע זו חוצה את החתך (S,V\S) אשר מכבד את A. T הכיל את הקבוצה A, ולכן מהבנייה T_1 מכיל את A ואת הצלע e'. הצלע e'' אינה ב-A, ומבחירתה היא גם אינה e', ולכן מבניית T_2 נובע כי העץ T_2 מכיל את $A \cup \{e'\}$.

3. מהנתון עבור e' נובע כי- $w(e'') \geq w(e')$. כיוון שכך נקבל כי $w(T_2) \leq w(T)$. ממינימאליות T נובע כי $w(T_2) = w(T)$ ולכן T_2 מינימאלי במשקלו.

<u>סעיף ב</u>

<u>תאור הרדוקציה</u>
יהי G=(V,E), w, e המופע עבור הבעיה.
שני פתרונות אפשריים:

<u>פתרון 1</u>
נשנה את פונקציית המשקולות כך:

$$w'(u,v)=\begin{cases} w(u,v) & (u,v)\neq e \\ \min_{e'\in E}\{w(e')\}-1 & (u,v)=e \end{cases}$$

נפתור את בעיית MST עבור המופע G=(V,E), w'.
נחזיר את העץ הפורש המינימאלי שהתקבל כפתרון לבעיית MST.

<u>פתרון 2</u>
נבנה גרף חדש G'=(V',E') כך:
V' תהיה קבוצת הקודקודים V, כך שקודקודי e יזוהו כקודקוד בודד.
$E'=E\setminus\{e\}$
נפתור את בעיית MST עבור המופע G'=(V',E'), w. ונקבל עץ פורש מינימאלי T'=(V',E$_{T'}$)
נחזיר את קבוצת הצלעות $E_{T'}\cup\{e\}$

<u>ניסוח טענה ראשית</u>
אלגוריתם הרדוקציה מחזיר פתרון אופטימאלי.

שאלה 3

<u>זמן ריצה</u>
$O(k\log k)$

<u>נימוק זמן ריצה</u>
כל פעולת איחוד מורכבת מחיפוש אחר מנהיג של שתי הקבוצות ואז העתקת המצביעים של איברי הקבוצה הקטנה להצביע למנהיג הקבוצה הגדולה. נסמן את עלות האיחוד ה-i ב- u. נחשב את עלותה של פעולת איחוד אחת $u_i=f_i+c_i$, כאשר f_i הוא עלות החיפוש אחר המנהיגים ו-c_i הוא עלות ההעתקות שמתבצעות באיחוד ה-i. עלינו לחשב את $\sum_{i\in\{1,\ldots,k\}}u_i=\sum_{i\in\{1,\ldots,k\}}f_i+\sum_{i\in\{1,\ldots,k\}}c_i$. נשים לב שלכ ל $1\le i\le k$, $f_1\in O(1)$, כיוון שמציאת מנהיגים והשוואתם לוקחת זמן קבוע. ולכן $\sum_{i\in\{1,\ldots,k\}}f_i=O(k)$.
כדי לחשב את $\sum_{i\in\{1,\ldots,k\}}c_i$, נשתמש בשיטת פיזור החיובים. נחשב כמה העתקות של מנהיג עוברות על איבר בודד. כיוון שעבור איזשהו איבר x, המנהיג אליו הוא מצביע משתנה רק אם x נמצא בקבוצה הקטנה יותר בזמן פעולת איחוד. ולכן כל פעם שהמנהיג של x משתנה, הוא הופך להיות חלק מקבוצה בגודל <u>לפחות פי 2</u> מהקבוצה הקודמת שלו.
היות והתחלנו מקבוצות בגודל 1, אזי לאחר k איחודים גודל הקבוצה המקסימאלית יהיה לכל היותר k.
ולכן כל איבר x יחליף לכל היותר $\log k$ מנהיגים! בנוסף מספר האיברים <u>השונים</u> שיכולים להשתתף ב k איחודים הוא לכל היותר k ולכן $\sum_{i\in\{1,\ldots,k\}}c_i\le k\log k$. ולכן,

$$\sum_{i\in\{1,\ldots,k\}}u_i=\sum_{i\in\{1,\ldots,k\}}f_i+\sum_{i\in\{1,\ldots,k\}}c_i\in O(k\log k).$$