אלגוריתמים – ממ"ן 12

שם : איתי אירמאי

תז : 204078224

תאריך הגשה : 14/12/2019

מייל : I.tai307@gmail.com

(הערה : בסקירת נוסח אצל רועי.)

1.

הגדרה כללית : יהא $P_{SV} = (s = u0,u1,u2,u3,\ldots,un = v)$ מסלול מסוים. קשת $e_i = (u_{i-1},u_i)$ תיקרא <u>שימושית במסלול P_{SV}</u> $\Leftrightarrow$ תת המסלול $(s,u1,u2,\ldots,u(i-1),ui)$ של P הינו מזערי. נבחין כי זהו מקרה פרטי של קשת שימושית.

א. יהא $P_{S,V} = (s = u0,u1,u2,u3,\ldots,un = v)$ המסלול הנדון, בו כל קשת $e_i = (u_i,u_{i+1})$ הינה שימושית.

נוכיח באינדוקציה כי אם כל הקשתות $ei, I = 0{:}\,K$ הינן שימושיות, אז תת המסלול $P_{s,ui}$ הינו מזערי. כאשר K=N, הטענה מוכחת.

<u>K=1</u> : אם $e1 = (s,u1)$ הינה שימושית, $P = (s,u1)$ מזערי על פי הגדרת משקל מסלול.

<u>K+1</u> : נחלק לשני מקרים.

① אם $e(k+1)$ שימושית במסלול $P_{s,v}$, אז על פי הגדרה $P_{s,u(k+1)}$ מזערי כנדרש.

② אחרת, $e(k+1)$ שימושית במסלול $P'_{s,u(k+1)} = (s,u1',u2',\ldots,uj',uk,u(k+1))$ אחר, שהינו **מזערי ומתלכד בקשת האחרונה** (לפחות) עם P (יכול להיות באורך שונה ממנו). על פי משפט, תת המסלול $P'_{s,uk}$ הינו מזערי, ועל פי הנחת האינדוקציה, $P_{s,uk}$ הינו מזערי $\Leftarrow$ באותו משקל. מבחינת משקלים :

$$w(P_{s,u(k+1)}) = w(P_{s,uk}) + w(e_{k+1}) = w(P'_{s,uk}) + w(e_{k+1}) = P(P'_{s,u(k+1)})$$

שני המסלולים P ו-P' עד u_{k+1} בעלי אותו משקל ו-P' מזערי, ולכן $P_{s,u(k+1)}$ מזערי כנדרש. ▌

ב. נניח בשלילה כי P הנדון מזערי. תהא $ei = (u_{i-1},u_i)$ אחת מהקשתות ב-P שאינה שימושית $\Leftarrow$ תת המסלול $P_{s,ui}$ אינו מזערי $\Leftarrow$ קיים מסלול $P'_{s,ui}$ שעבורו $w(P'_{s,ui}) < w(P_{s,ui})$. ניצור את המסלול הבא, $P''_{s,v} = (P'_{s,ui},P_{ui,v})$ ע"י איחוי P' עם שארית P.

$$\mathbf{w(P''_{s,v})} = w(P'_{s,ui}) + w(P_{ui,v}) < w(P_{s,ui}) + w(P_{ui,v}) = \mathbf{w(P_{s,v})}$$

המסלול המאוחה קצר יותר מ-P, בסתירה למזעריותו $\Leftarrow$ P אינו מזערי. ▌

ג. נניח בשלילה כי במסלול הנדון P (הכמעט מזערי) קיימות שתי קשתות לא שימושיות (לפחות) ; נקרא לשתיים הראשונות מביניהן $ei = (u_{i-1},u_i)$ ו-ej. קיימת לפחות אחת כי הוכחנו בסעיף א שמסלול עם 0 קשתות לא שימושיות הינו מזערי.

עקב קיום מסלול מ-s לכל קודקוד, קיים מסלול מזערי בין s ל-ui – נקרא לו P1 ; וכן מסלול מזערי בין s ל-uj – נקרא לו P2. שניהם בהכרח שונים (קלים יותר) מתתי המסלול של P לכל אחד משני הקודקודים (שאינם מזעריים עקב אי שימושיות הקשתות) ; אך יכולים להתלכד זה עם זה (קרי, P1 תת מסלול של P2). נגדיר את המסלולים הבאים :

① $P'_{s,v} = (P1_{s,ui},P_{ui,uj},P_{uj,v})$.

② $P''_{s,v} = (P2_{s,uj},P_{uj,v})$.

תת המסלול עד uj של P' אינו מזערי, מכיוון שהוא מכיל את ej הלא שימושית – <u>כבד לעומת P2</u>.

$$\mathbf{w(P_{s,v})} = w(P_{s,ui}) + w(P_{ui,uj}) + w(P_{uj,v}) > w(P1_{s,ui}) + w(P_{ui,uj}) + w(P_{uj,v}) = \mathbf{w(P'_{s,v})}$$

$\mathbf{w(P'_{s,v})} = w(P1_{s,ui}) + w(P_{ui,uj}) + w(P_{uj,v}) > w(P2_{s,uj}) + w(P_{uj,v}) = \mathbf{w(P''_{s,v})}$

על כן, קיבלנו $w(P_{s,v}) > w(P'_{s,v}) > w(P''_{s,v})$, בסתירה לכך ש-P הינו כמעט מזערי

⇐ קיימת בדיוק קשת לא שימושית אחת ב-P. ▌

ד. נשנה במקצת את הסימון בהתאם להגדרות הנ"ל : $ei = (u_{i-1},u_i)$ הינה הקשת הלא שימושית.

על פי הוכחת האינדוקציה בסעיף א, הרישא או מסלול $P_{s,ui-1}$ הינו מזערי, כי כל הקשתות עד e_{i-1} כולל הינן שימושיות.

נניח בשלילה כי הסיפא $P_{ui,v}$ אינו מזערי. יהא $P'_{s,ui}$ מסלול מזערי מ-s ל-ui. על פי הגדרה, כל קשתותיו שימושיות (במסלול P') ; כמו כן, כל קשתות $P_{ui,v}$ שימושיות כי אינה מכילה את הקשת הלא שימושית ei (לפי סעיף ג). נתבונן במסלול $P''_{s,v} = (P'_{s,ui},P_{ui,v})$: זהו מסלול בין s ל-v, שהינו מזערי כיוון שכל קשתותיו שימושיות. לפי משפט, כל תת מסלול של P'' צ"ל מזערי, ספציפית $P_{ui,v}$, בסתירה להנחה. ⇐ גם הסיפא הינו מזערי. ▌

ה. תיאור : כמעט העתק הטכניקה מתרגיל 11.4 ; נבצע רדוקציה מבעיית איתור מסלול כמעט מזערי ב-G לאיתור מסלול מזערי בגרף דו שכבתי G' עם קישור דרך קשתות לא שימושיות (ראו מטה).

ממיר קלט : נחלק את כל הקשתות לשימושיות ולא שימושיות ע"י זיהוי **כל** המסלולים המזעריים ב-G (הפעלת Dijkstra) ; ניצור G' כדלקמן :

$G'=(\{V1 \cup V2\},\{E1 \cup E2 \cup E3\})$; $V1 = \{v.1 \mid v \in V\}$; $V2 = \{v.2 \mid v \in V\}$;

$E1 = \{(u.1,v.1) \mid Inefficient[(u,v)] = 0\}$; $E2 = \{(u.2,v.2) \mid Inefficient[(u,v)] = 0\}$;

$E3 = \{(u.1,v.2) \mid Inefficient[(u,v)] = 1\}$

קופסא שחורה : הפעלת Dijkstra מקודקוד s.1 לקודקוד t.2.

ממיר פלט : אם לא קיים מסלול מ-s.1 ל-t.2 אז להחזיר שלא קיים מסלול כמעט מזערי בין s ל-t (מצב כזה אפשרי כאשר כל הקשתות הלא שימושיות מובילות החוצה ממסלול s-t). אחרת להחזיר מסלול מזערי מ-s.1 ל-t.2 כאשר לכל קודקוד v.2 / v.1 במסלול נחזיר את $v \in V$ המקביל.

פירוט :

קלט : גרף מכוון G=(V,E) בצורת רשימת סמיכויות.

פקודות :

1) צור מערך בוליאני בגודל |E| / טבלת גיבוב בשם "Inefficient" עם מפתחות מ-E וערך בוליאני 1 לכל קשת. גישה ישירה ב-O(1) ולכל הערכים ב-O(E).

2) הרץ Dijkstra מקודקוד s עם התאמת "איתור מסלולים מרובים" :

- האלגוריתם ישמור רשימת אבות מינימאליים לכל קודקוד, בנוסף למשקל המסלול המזערי.

- הוסף תנאי (בצמוד ל-RELAX) : בהינתן צומת v חדש לבדיקה ו $u \notin S$ שכן שלו ;

אם $dist(v) + weight(u,v) = dist(u)$, אזי הוסף את צומת v לרשימת האבות המיני' של u.

-* בתוך תנאי RELAX : אם $dist(v) + weight(u,v) < dist(u)$, אזי צור רשימת אבות מיני' חדשה לצומת u והוסף אליה את v.

3) צור גרף מכוון G' (ייצוג רשימה) כדלקמן :

$G'=(\{V1 \cup V2\},\{E1 \cup E2 \cup E3\})$; $V1 = \{v.1 \mid v \in V\}$; $V2 = \{v.2 \mid v \in V\}$;

$E1 = E2 = E3 = \emptyset$

4)** עבור כל קודקוד ב-V, רוץ על רשימת האבות שלו בפלט 2 (Dijkstra) ;

בעבור קודקוד v ואב u, סמן 0 בטבלת הגיבוב Inefficient על הקשת (u,v).

5) עבור על כל הקשתות ב-Inefficient וצור קשת בין שתי השכבות אם מסומן 1, או קשת תוך שכבתית אם מסומן 0, <u>עם אותו משקל קשת כמו במקור</u>, כלומר:

$E1 = \{(u.1,v.1) \mid Inefficient[(u,v)] = 0\}$

$E2 = \{(u.2,v.2) \mid Inefficient[(u,v)] = 0\}$

$E3 = \{(u.1,v.2) \mid Inefficient[(u,v)] = 1\}$

6) הרץ Dijkstra על גרף G' מקודקוד s.1, עם סימון אב יחיד בלבד.

7) השתמש בפלט 6 על מנת למצוא מסלול מ-s.1 ל-t.2 באמצעות סימוני האבות (כלומר, בסדר הפוך – התחל מ-t.2 וחזור עד ל-s.1), אם הוא קיים. <u>החזר מסלול זה כמסלול כמעט מזערי של G ע"י הסרת אינדקס 1 / 2 מכל קודקוד</u>.

* החלפת רשימה קיימת בחדשה עקב משקל נמוך יותר תלווה בפעולת ניקוי או תושאר לאיסוף אשפה. ניקוי של כל הרשימות הקיימות יהיה בסה"כ O(E).

** חייבים לרוץ על כל הקודקודים ולא רק על המסלול $s \to t$ כי גם קשתות שימושיות מחוץ למסלול המזערי יכולות לשמש מסלול כמעט מזערי, אם למשל קיימים רק שני מסלולים נפרדים לגמרי מ-s ל-t. הזמן הוא O(E) בהתאמה לייחודיות מופעי הקשתות ופשטות הגישה.

<u>סיבוכיות</u>: זמן - עבור הרצת Dijkstra ‏O(|E|*log|V|) (גם עם x2 קודקודים וקשתות בפקודה 5).

- טבלת הגיבוב אוכלת קצת מקום, אבל גישה ועדכון O(1), יצירה O(E).

- התאמת "מסלולים מרובים" אינה משנה את חסם הסיבוכיות – הוספה לרשימה היא O(1), ניקוי הרשימות כאמור יעלה O(E) בסה"כ, ועדיין בודקים כל קשת פעם אחת בדיוק + איתור קודקוד מינימאלי לאיטרציה ללא שינוי.

- ריצה על תת קבוצות האבות / קשתות מינימאליות זניחה, כמובן; יחד עם טבלת הגיבוב המבטיחה שזמן יצירת E1, E2, E3 יהיה מינימאלי – O(E).

<u>הוכחה</u>:

טענת עזר: קיים מסלול במשקל X בין s ל-t שעובר בקשת לא שימושית יחידה בגרף G $\Leftrightarrow$

קיים מסלול במשקל X בין s1 ל-t2 בגרף G'.

הוכחת טענת עזר: ① יהא $P = (s=u_0,u_1,u_2,\ldots,u_k,u_{k+1},\ldots,u_n=t)$ מסלול העובר בקשתות שימושיות בלבד פרט לקשת (u_k,u_{k+1}) בגרף G. אזי מסלול

$P' = (s.1,u_1.1,u_2.1,\ldots u_k.1,u_{k+1}.2,\ldots,u_n.2=t.2)$ הינו חוקי ב-G' (הקשת בין שכבה 1 ל-2 קיימת כי

(u_k,u_{k+1}) אינה שימושית, ושאר הקשתות בתוך שכבה 1 ו-2 קיימות כי הן שימושיות), ובעל אותו משקל.

② יהא $P' = (s.1,u_1.1,u_2.1,\ldots u_k.1,u_{k+1}.2,\ldots,u_n.2=t.2)$ מסלול בגרף G'. אזי מסלול

$P = (s=u_0,u_1,u_2,\ldots,u_k,u_{k+1},\ldots,u_n=t)$ הינו חוקי ב-G (על פי הגדרת הגרף G', כל קשת בין שני קודקודים קיימת בגרף המקורי ללא אינדקס שכבה), בעל אותו משקל, ומכיל קשת לא שימושית אחת בלבד (u_k,u_{k+1}) עקב הגדרת החלוקה E1 / E2 / E3 (קשתות בין שכבות אינן שימושיות).

משפט: המסלול המזערי $P^* = (s.1,u_1.1,u_2.1,\ldots u_k.1,u_{k+1}.2,\ldots,u_n.2=t.2)$ בגרף G' המוחזר ע"י האלגוריתם מתאים למסלול כמעט מזערי בין s ל-t, שהינו $P = (s=u_0,u_1,u_2,\ldots,u_k,u_{k+1},\ldots,u_n=t)$; אם לא הוחזר מסלול, לא קיים מסלול כמעט מזערי בין s ל-t.

הוכחת משפט: ① אם הוחזר מסלול P* כאמור, על פי ט"ע קיים P בעל אותו משקל. אם בשלילה קיים מסלול קצר יותר P' ב-G העובר בקשת לא שימושית יחידה,

לפי ט"ע $\exists P'^*$ ב-G' בו $w(P'^*) < w(P^*)$, וזאת בסתירה לנכונות Dij.

$\Leftarrow$ P הינו בעל המשקל המזערי מבין המסלולים בעלי קשת לא שימושית אחת. בנוסף לכך, הוכחנו בסעיף א שכל המסלולים בהם <u>כל הקשתות שימושיות</u> הינם מזעריים (כלומר בעלי אותו משקל <u>W1</u>), ובסעיף ב

+ ג שכל המסלולים בהם <u>יותר מקשת לא שימושית אחת</u> אינם מזעריים ואינם כמעט מזעריים (כלומר <u>משקליהם $\geq$ W3</u>) ;

⇐ לכן, <u>מסלול P</u> בעל קשת לא שימושית אחת ומשקל מזערי בקטגוריה חייב להיות <u>כמעט מזערי (W2)</u>.

② אם לא קיים מסלול P^* מ-s.1 לt.2 ב-G׳, נניח בשלילה שקיים מסלול כמעט מזערי P׳ בגרף G ⇐ מסלול כמעט מזערי הינו בעל קשת לא שימושית יחידה על פי סעיף ג, ולכן בהכרח קיים מסלול מקביל P'^* באותו משקל ב-G׳ על פי ט״ע ⇐ סתירה לאי קיום P^*, וסיימנו. ▌

.2

<u>תיאור</u>: נשמיט את הקשת e^*, נריץ BFS / DFS פעמיים כדי לקבל שני רכיבי קשירות שנוצרו כתוצאה מכך, ונחפש את הקשת הקצרה ביניהם.

<u>פירוט</u>:

קלט:

$G = (V,E)$ גרף לא מכוון, T עפ"מ של G, קשת e^* כלשהי ב-G ו-T.

פקודות:

1) צור טבלת גיבוב ריקה בשם "קשירות1" [או 'קבוצה', set בפייתון – בדיקת קיום איבר ב-$O(1)$, או לצורך הפשטה מערך בוליאני מאותחל באפסים בגודל $|V|$];

רשימה ריקה "קשירות2";

קשת "חיבור מינימאלי" ריקה (עם משקל אינסוף).

2) השמט את צלע e^* מ-T ו-G.

3) הרץ DFS על **T** החל מצומת כלשהו;

הוסף כל צומת שבוקר לטבלה קשירות1.

4) הרץ DFS על **T** החל מצומת כלשהו שטרם בוקר;

הוסף כל צומת שבוקר בשלב זה לרשימה קשירות2.

5)* עבור כל צומת u בקשירות2, בצע:

5.1)* עבור כל קשת (u,v) הקיימת ברשימת הקשתות של u ב-G, בצע:

5.1.1)* אם v נמצא בקשירות1 ומשקל (u,v) קטן ממשקל חיבור_מינימאלי, בצע:

5.1.1.1)* עדכן את חיבור_מינימאלי לקשת (u,v).

6) הוסף את חיבור_מינימאלי ל-T והחזר את T.

* או בקיצור, "אתר את הקשת המזערית מקשירות2 לקשירות1".

<u>סיבוכיות</u>: $O(|E|)$ זמן.

- DFS בעיקרון $O(|V| + |E|)$ אבל ה-V זניח (שקול) כי מריצים על עץ.

- מעבר על קשתות כל הקודקודים (ברכיב קשירות) ובדיקת קיום בטבלה – $O(|E|)$ לכל היותר.

- השמטות והוספת קשת – $O(1)$.

<u>הוכחה</u>: (הערה: ראיתי את המשפט בספר, על קשת מזערית בין שני רכיבי קשירות; לא השתמשתי בו בהוכחה כי הוא מניח עלות שונה לכל הקשתות, והנחה זו אינה תקפה כאן. די מאריך את הדרך.)

נסמן $T' = T \setminus \{e^*\}$. נסמן $T1 = (V1,E1)$, $T2 = (V2,E2)$ כאשר $T' \equiv T1 \cup T2$ בחלוקה

$T1 \cap T2 = \emptyset$; כלומר, חלוקה של כל הקודקודים והקשתות מהעץ המפורק לשני תתי גרפים, אשר כל אחד מכיל רק קודקודים ביניהם קיימות קשתות ואת הקשתות המשויכות להם (קרי, <u>רכיבי קשירות</u>).

נחלק ל-3 טענות עזר.

① <u>קיימת</u> חלוקה כזאת – כלומר, ההפרדה מחזירה <u>בדיוק שני</u> רכיבי קשירות.

② T1 ו-T2 הינם עפ״מ על V1 ו-V2 ב-G (וגם G׳) בהתאמה, תוך התעלמות מקשתות חיצוניות.

כלומר, T1 עפ״מ על $G'' = (V1,\{(u,v) \mid \forall(u,v) \in E \ \&\ u,v \in V1\})$ ו-T2 בדומה לו.

③ T״ הנבנה משני רכיבי T1 + T2 והקשת הקלה ביותר (e׳) בין T1 ל-T2 הינו עפ״מ של G׳.

בהינתן הנ״ל, <u>האלגוריתם מבטיח זיהוי</u> שני תתי העצים ע״י שימוש ב-DFS (הגרף אינו מכוון אז לא נוצרים רכיבי קשירות פיקטיביים), ומאתר את <u>קשת החיבור הקלה</u> e׳ מסעיף 3 בחיפוש <u>מקיף</u>, כנדרש – האלגוריתם עובד.

<u>הוכחת ט״ע1</u>: נגדיר את הקשת המושמטת $(u,v) = e^*$. מכיוון ש-T הוא עץ פורש, על פי משפט קיים מסלול פשוט <u>יחיד</u> בין כל שני קודקודים ב-T, ספציפית בין u ל-v (והוא $P = (e^*)$). לאחר השמטת e^* אין מסלול בין u ל-v, ולכן הם שייכים לשני רכיבי קשירות שונים ⇐ קיימים לפחות שני רכיבי קשירות. נניח בשלילה שקיימים יותר משניים ⇐ קיים $x \in V$, שאינו ברכיבי הקשירות של u,v ⇐ לא קיים מסלול בין u, v וכל הקודקודים ברכיבי הקשירות שלהם ל-x ב-T׳; הוספת קשת e^* חזרה ל-T׳ תיצור את T מחדש ותחבר בין u ל-v, אך בהיעדר מסלול ל-x מכל אחד מהם עדיין לא יהיה מסלול מ-u או v ל-x, בסתירה לקשירות עץ פורש T.

<u>הוכחת ט״ע2</u>: חילקנו את כל הקודקודים לשני רכיבי קשירות T1 ו-T2 – על פי הגדרה הם קשירים. כמו כן, $T1, T2 \subseteq T' \subseteq T$ (הקודקודים והקשתות מוכלים ב-T), ו-T היה חסר מעגלים, על כן כל תת גרף גם הינו חסר מעגלים (אחרת אותו מעגל C בתת גרף בהכרח היה קיים בגרף המלא עם אותם קודקודים) ⇐ על פי משפט, T1 ו-T2 הינם עצים פורשים ברכיבים.

לגבי המזעריות – נניח בשלילה כי סך המשקלים ב-T1 או T2 אינו מזערי; קיים עץ פורש T1׳ או T2׳ (ההוכחה סימטרית) המקיים

$$\sum w(T1') \le \sum w(T1) \text{ OR } \sum w(T2') \le \sum w(T2)$$

בנוסף, כל הקשתות של T מחולקות בין T1, T2 ו-e^*.

$$\Rightarrow \sum w(T) = \sum w(T1) + \sum w(T2) + w(e^*) >= \sum w(T1') + \sum w(T2) + w(e^*)$$

החיבור של T1׳, T2 ו-e^* הינו עץ פורש ב-G (קשיר וללא מעגלים – איחוד טריוויאלי מהגדרותיהם), ובעל משקל קטן יותר מ-T בסתירה למזעריות T.

⇐ סך המשקלים ב-T1, T2 מזערי.

<u>הוכחת ט״ע3</u>: תהא $(e' = (u',v') \mid u' \in V1, v' \in V2)$ הקשת המינימאלית המחברת בין V1 ל-V2 ב-G׳, כמתואר – זו קיימת כי נתון ש-G׳ קשיר. ניתן ליישם כאן את המשפט ״״״ בגרף לא מכוון G עם עץ פורש T וקשת e שאינה ב-T, הגרף H הבנוי מקשתות T + e מכיל מעגל, וניתן להסיר כל קשת אחרת e׳ ב-H ונקבל עץ פורש אחר T׳ ״״״: הוספת הקשת e׳ ל-T סוגרת מעגל, ולאחר השמטת e^* מתקבל אותו T״ שהיה מתקבל כתוצאה מהיפוך הסדר כפי שבוצע באלגוריתם (השמטת או הוספת שתי קשתות שונות הינן פעולות בלתי תלויות).

⇐ T״ זה הינו עץ פורש ב-G, וגם ב-G׳ כי אינו מכיל את e^*. צ״ל שהינו מזערי.

נניח בשלילה שכל עץ מזערי ב-G׳ מכיל לפחות שתי קשתות בין V1 ל-V2 – בעיקרון הדרך היחידה לפסול את T״, כי T1 ו-T2 נפרדים (כלומר, אפס חיבורים מביאים לאי קשירות).

נקרא לשתיים מהקשתות $e1' = (u1,v1), e2' = (u2,v2)$ כאשר $u1,u2 \in V1$ ו $v1,v2 \in V2$.

ב-T, הקשתות $f1 = (u2,u3)$ ו $f2 = (v2,v3)$ חיברו את שני הקודקודים u2,v2. בהכרח,

$$w(e2') + w(f1), w(e2') + w(f2) \ge w(f1) + w(f2) \Rightarrow w(e2') \ge w(f1), w(f2)$$

זאת בעקבות מזעריות T על G והעובדה ששלושת השילובים הנ״ל מייצרים מסלול יחיד לשני הקודקודים (דרך $u3 \to u2 \to v2$ או $u2 \to v2 \to v3$ או $u2 \to \dots e^* \dots \to v2$), עם פוטנציאל לעץ פורש אחרי הוספת שאר הקשתות מ-T. כמובן, f1 ו-f2 קיימות גם ב-G׳ (כל אחת מוכלת ב T1 / T2 בהתאמה).

אם כן, נחליף את הקשת e2׳ ב-f1 או f2, וקיבלנו עץ פורש על G׳ בעל משקל ≥ עץ מזערי ב-G׳ עם קשת יחידה בין V1 ל-V2 (קרי e1׳), בסתירה להנחה.

⇐ קיים עץ מזערי ב-G׳ המכיל בדיוק קשת אחת בין V1 ל-V2 (ב׳מבנה׳ של T״).

משקל העץ T״ כפי שהוגדר לעיל: $w(T'') = w(T1) + w(T2) + w(e')$

עקב היות T״ עץ פורש מהמשפט, <u>מזעריות T1 ו-T2</u> שהוכחה בסעיף 2, מזעריות הקשת <u>e׳</u> שנבחרה והוכחת <u>קיום עפ״מ</u> על G׳ עם <u>המבנה</u> של T״, נובע כי T״ מזערי מבין העצים הפורשים בעלי מבנה דומה ⇐ <u>T״ הינו עפ״מ של G׳</u>, וסיימנו. ▮

3.

אם מותרות חזרות, דוגמת נגד פשוטה:

$$\varphi = (x1 \mid x1 \mid x2) \;\&\; (x1 \mid x1 \mid x3) \;\&\; (\overline{x1} \mid \overline{x1} \mid \overline{x1})$$

x1 מופיע 4 פעמים בחיוב ו-3 בשלילה, ולכן מושם כ-T, שאינו מספק את הפסוקית השלישית.

לעומת זאת, ההשמה **X1=F, X2 = T, X3 = T** מספקת.

ללא חזרות צריך לוגיקה יותר מורכבת:

$$\varphi = (\mathbf{x1} \mid x2 \mid \overline{x3}) \;\&\; (x2 \mid x3 \mid \overline{\mathbf{x1}}) \;\&\; (\mathbf{x1} \mid x3 \mid \overline{x2}) \;\&\; (\mathbf{x1} \mid x4 \mid x5) \;\&\; (x2 \mid x6 \mid x7) \;\&$$

$$\& \; (x3 \mid x8 \mid x9) \;\&\; (\mathbf{x1} \mid x10 \mid x11) \;\&\; (x2 \mid x12 \mid x13) \;\&\; (x3 \mid x14 \mid x15) \;\&\; (\overline{\mathbf{x1}} \mid \overline{x2} \mid \overline{x3})$$

x1 / x2 / x3 מופיעים 4 פעמים בחיוב ו-2 בשלילה כל אחד, ולכן מושמים כ-T (אפילו אם יש במקרה ספירה מחדש אחרי כל השמה, בהנחה והלולאה עולה בסדר האינדקסים), וזה לא מספק את הפסוקית האחרונה.

לעומת זאת, ההשמה **X1=F, X2 = F, X3 = F** ו-**X4:15 = T** למשל מספקת.

4.

<u>משפט</u>: לכל עץ בינארי מלא קיימת סדרת שכיחויות $f1, f2, \ldots, fn$ כך שקידוד הופמן של הסדרה מקביל לעץ במבנה זה.

<u>הוכחה</u>: יהא T עץ בינארי מלא בעל n עלים בעומק k. עבור כל צומת בעץ זה מלבד השורש, נסמן את <u>מיקום הצומת</u> בעץ לפי <u>קוד התחיליות שלו בבסיס 2</u>; הווה אומר הבן השמאלי של השורש בעל מיקום 0, הבן הימני של השורש בעל מיקום 1, הבן הימני של הבן השמאלי של השורש בעל מיקום 01 וכד׳.

עבור כל עלה I ב-T, הנמצא בעומק $d(i)$ בעץ, נגדיר התאמה <u>לסדרת שכיחויות f1..fn</u>, כך ש:

$\mathbf{fi = c*2^{k-d(i)}}$, כאשר $\mathbf{c} \in \mathbb{N}$ כלשהו. לשם הנוחות, מספור העלים יוגדר קודם לפי עומקם, ואח״כ לפי אינדקס המיקום. לדוגמא, לכל עלה ברמת העומק האחרונה תוגדר שכיחות 1, ולעלה ברמה מעליו שכיחות 2. לכל ערך שכיחות <u>יוצמד ערך המיקום</u>.

נסתכל על ריצת אלגוריתם הופמן מסוימת H על סדרת השכיחויות: נגדיר שבכל נקודה בריצה זו, אם למספר צמתים קיים <u>ערך שכיחות זהה</u>, הם ימוינו על פי <u>המיקום</u>; וכאשר מאוחדים שני צמתים בעלי אב משותף (קרי, המיקום זהה פרט לספרה האחרונה – נראה בהמשך כי זה תמיד המקרה), צומת האב שייווצר כתוצאה מהאיחוד יקבל אינדקס מיקום זהה לבנים תוך <u>השמטת הספרה</u> האחרונה. אם ערכי המיקום של העלים וכל אב הנוצר בריצה H תואמים למבנה T, אז H בנה את T.

טענה: ריצה H תבנה עץ במבנה של T.

הוכחה <u>באינדוקציה על עומק</u> העץ T. נתחיל מהמקרים $K=1, K=2$ (שכן לאלגוריתם הופמן קיימת החרגה בעומק זה).

<u>$K=1$</u>: ל-T קיים עלה אחד. $F1 = c$, אז H מחזיר עלה אחד כנדרש.

<u>$K=2$</u>: ל-T קיים שורש ומתחתיו שני עלים. $F1 = F2 = c$, אז H מאחד את F1 ו-F2 עבור כל ערך C, יוצר להם צומת אב בעל שכיחות 2C ומסיים, בהתאם למבנה.

<u>K</u>: נניח שלכל עץ בינארי מלא T׳ בגובה K, הרצה H על סדרת שכיחויות $fi = c * 2^{k-d(i)}$ המותאמת למבנה ולכל C טבעי, מחזירה עץ שמבנהו זהה ל-T׳.

<u>$K+1$</u>: יהא T עב״מ כלשהו בעומק $K+1$. נתאים לעליו סדרת שכיחויות fi כמצוין לכל C, ונתבונן בריצת H על סדרה זו (או סדרות אלו).

ברמת העומק $K+1$ קיימים רק עלים על פי הגדרת עומק. לכל צומת בעב״מ חייבים להיות 0 או 2 בנים ⇐ לכל עלה קיים צומת אח ⇐ ב-$K+1$ מספר העלים זוגי. על פי הגדרת סדרת השכיחויות, fi של העלים בעומק זה אחיד (= C), ושל כל העלים האחרים < C (מוכפל בחזקה של 2). כל אב שנוצר מאיחוד בין שני עלים מרמה $K+1$ יהיה בעל שכיחות 2C על פי הגדרת האלגוריתם, הזהה לשכיחות של עלים ברמה K, ועל כן ימוין בין הצמתים של רמת עומק K במיקום המצופה ב-T. בנוסף, על פי הגדרת H, כל שני אחים ברמת העומק ימוקמו בסמיכות זה לזה.

בעקבות הנ״ל, כל צעד מבין X צעדים בשלב ה׳ראשון׳ של H מאחד בין שני עלים ב-$K+1$ המתאימים זה לזה, ויוצר להם אב חדש wj׳ בעל שכיחות 2C ומיקום תואם לעץ. אין איחוד בין שתי רמות עומק שונות מכיוון שמספר העלים ברמה $K+1$ זוגי, איחוד מופעל על שני איברים ויוצר איבר שאינו רלוונטי ל-X השלבים. (כלומר, לפני צעד קיימים 2L עלים בשכיחות 1 ואחריו נותרים $2L – 2$, עד ש-$L = 1$.)

לאחר הרצת X צעדים של H, קיבלנו סדרת שכיחויות חדשה: $fi' = C * 2^{(k+1) – d(i)} = 2C * 2^{k – d(i)}$, התואמת לתת העץ של T עד עומק K (עם מיקומים נכונים של ה׳עלים׳, תוך מחיקת זנבות $K+1$ מאלפאבית S).

נגדיר $C' := 2*C$. לפי הנחת האינדוקציה, (המשך) הפעלת H על סדרת שכיחויות fi׳ עם C׳ אכן יוצרת עץ T׳ בעומק K במבנה המצופה; יחד עם העלים מרמה $K+1$, להם הוגדר מיקום נכון בשלב הראשון, קיבלנו את העץ T מסדרת השכיחויות fi. ▌