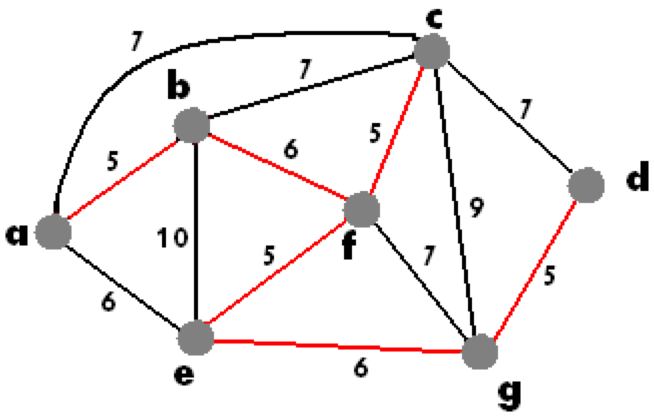
# בעיית העץ הפורש המינימאלי (MST)

נניח שקיימת קבוצת איים שאנו מעוניינים לקשר ביניהם על ידי גשרים, כך שיהיה ניתן לנסוע מאי אחד לכל אי אחר מקבוצה זו. בנוסף, נניח כי הממשלה רוצה להוציא את הסכום המינימאלי האפשרי על פרויקט זה. המהנדסים מסוגלים לחשב את העלות של בניית גשר בין כל זוג איים. מערכת הגשרים שתאפשר לנסוע מכל אי לכל אי ושעלות הקמתה מינימאלית, הינה עץ פורש מינימאלי.
בדוגמא שלפניכם יש גרף ו-MST שלו:




- בתרגול זה נעסוק רק בגרפים פשוטים, לא מכוונים, אך ממושקלים.

<u>הגדרות:</u>

**גרף פשוט-** גרף $G$ הוא פשוט אם אין בו צלעות עצמיות (צלע מקודקוד לעצמו) או צלעות כפולות (יותר מצלע אחת בין זוג קודקודים).

**מסלול** – מסלול בגרף $G=(V,E)$ הוא סידרת קודקודים $<v_0,...,v_k>$, כאשר לכל $0 \le i < k$ מתקיים $(v_i, v_{i+1}) \in E$.

**מסלול פשוט** – מסלול $<v_0,...,v_k>$, הוא פשוט אם כל קדקודיו שונים זמ"ז, כלומר לכל $0 \le i < j \le k$ מתקיים $v_i \ne v_j$.

**מעגל -** מסלול $<v_0,...,v_k>$, כך ש-$v_0 = v_k$.

**עץ פורש** – עץ פורש של גרף קשיר $G=(V,E)$ הוא תת-גרף $T=(V,E_T)$ של G שמכיל את כל קדקודי G וחלק מצלעותיו (כלומר $E_T \subseteq E$), והוא מקיים:
1. $T$ חסר מעגלים.
2. $T$ קשיר (כלומר לכל שני קודקודים קיים מסלול המקשר ביניהם).

משפט 1:

יהי $G$ גרף (פשוט ולא מכוון). התנאים הבאים שקולים זה לזה:

1. $G$ קשיר וחסר מעגלים,
2. $G$ חסר מעגלים ו- $|E|=|V|-1$,
3. $G$ קשיר ו- $|E|=|V|-1$,
4. יש ב-$G$ מסלול פשוט יחיד בין כל זוג צמתים.

כדי להוכיח שתת הגרף $T=(V,E_T)$ של הגרף $G=(V,E)$ הוא עץ פורש של $G$, מספיק להראות שאחד מהתנאים הנ"ל מתקיים.

משפט 2:

יהי $G$ גרף, $T=(V,E_T)$ עץ פורש של $G$ ו- $e=(u,v)\in E\backslash E_T$ . הגרף $H=(V,E_T\cup\{e\})$ מכיל מעגל ולכל צלע $e'=(a,b)\in E_T$ במעגל, הגרף $T'=(V,E_T\cup\{e\}\setminus\{e'\})$ הוא עץ פורש של $G$.

(הוכח בכיתה)

הוכחה:

$T$ קשיר (כי הוא עץ פורש), ולכן $H$ (אשר נוצר כתוצאה מהוספת $e$ ל-$T$) קשיר. $H$ מכיל $|V|$ צלעות, לכן הוא אינו עץ ועפ"י תנאי 1 במשפט 1, ב- $G$ יש מעגל.

$T'$ עץ פורש:

1. ב-$T'$ ישנן בדיוק $|V|-1$ צלעות, שכן ב- $E_T$ היו $|V|-1$ והחלפנו צלע אחת.
2. $T'$ קשיר: ע"מ להוכיח כי $T'$ קשיר נראה כי עבור כל זוג קודקודים $s,t\in V$, קיים מסלול מ- $s$ ל- $t$ ב-$T'$.
   אם קיים מסלול מ- $s$ ל- $t$ ב- $T$ שלא כולל את הצלע $e'=(a,b)$, אזי אותו מסלול בדיוק קיים גם ב-$T'$.
   אחרת המסלול הוא $(s,...,a,b,...,t)$. ניזכר כי קיים מסלול $(u,\dots,a,b,\dots,v)$ ב- $T$ (כך בחרנו את $e'$). אזי המסלול $(s,....,a,....,u,v,.....,b,.....,t)$ כולו ב-$T'$ .
   (שימו לב ש-$T'$ קשיר לא מקביל לתנאי 4 במשפט 1!)

מתנאי 3 במשפט 1 ניתן להסיק ש-$T'$ עץ פורש של $G$ .

**עלות של עץ פורש** – באופן כללי, ניתן לבנות עצים פורשים רבים עבור גרף $G$. בהינתן גרף $G=(V,E)$ ופונקצית משקל $w:E\to R$ , נגדיר **עלות** של עץ פורש $T$ להיות סך המשקלות של כל צלעותיו, כלומר $w(T)=\sum_{e\in T}w(e)$ .

**עץ פורש מינימאלי (MST)-** של גרף $G$ הוא עץ פורש שעלותו מינימאלית מבין כל עלויות העצים הפורשים את $G$, כלומר נבקש למצוא עץ פורש $T$ כך ש : $w(T)=\min_{\substack{T'\text{ is a}\\ \text{spannig tree}}}\{w(T')\}$ .

(הערה- עץ פורש מינימאלי אינו בהכרח יחיד. יתכנו כמה כאלו, אבל לכולם, כמובן, אותו משקל.)

## **<u>שאלה 1:</u>**

נתון $G=(V,E)$ גרף לא מכוון ו- $w:E\rightarrow R$, פונקצית משקל על קשתות הגרף. יהי $T$ עץ פורש מינימאלי של $G$ תחת פונקצית המשקל $w$. נגדיר את $w':E\rightarrow R$ להיות פונקצית משקל באופן הבא: $w'(e)=w(e)+c$, כאשר $c\in R$ קבוע כלשהו. האם $T$ הוא עץ פורש מינימאלי של $G$ תחת פונקצית המשקל $w'$?

**<u>פתרון:</u>**

נוכיח כי $T$ הוא עץ פורש מינימאלי של $G$ תחת פונקצית המשקל $w'$. נניח בשלילה כי קיים עץ פורש של $G$, $T'$ כך ש- $\sum_{e\in T}w'(e)<\sum_{e\in T'}w'(e)$. ממשפט 1, אנו יודעים כי מספר הצלעות ב- $T$ ו- $T'$ הוא בדיוק $|V|-1$ ולכן,

$$\sum_{e\in T}w(e)+c(|V|-1)=\sum_{e\in T}w'(e)<\sum_{e\in T'}w'(e)=\sum_{e\in T'}w(e)+c(|V|-1)$$

לכן $\sum_{e\in T}w'(e)<\sum_{e\in T'}w(e)$ בסתירה לכך ש- $T$ הוא עץ פורש מינימאלי של $G$ תחת פונקצית המשקל $w$

**הערה**: שימו לב כי מה שהראינו הוא שאם מוסיפים קבוע למשקל של כל צלע, העץ הפורש נשאר ללא שינוי, זאת מכיוון שבסה"כ הוספנו עלות קבועה, עקב מספר קבוע של צלעות בכל עץ פורש - $|V|-1$.

## **<u>שאלה 2:</u>**

נתון $G=(V,E)$ גרף לא מכוון ו- $w:E\rightarrow R$, פונקצית משקל על קשתות הגרף. מצא עץ פורש מקסימאלי של $G$.

**<u>פתרון:</u>**

נבצע רדוקציה לבעיית העץ הפורש <u>המינימאלי</u>. נגדיר פונקציה משקל $w':E\rightarrow R$ באופן הבא, $w'(e)=-w(e)$. יהי $T$ עץ פורש מינימאלי של $G$ תחת פונקציה המשקל $w'$. נוכיח כי $T$ הוא עץ פורש <u>מקסימאלי</u> של $G$ תחת פונקצית המשקל $w$.

שוב, נניח בשלילה כי קיים $T'$ עץ פורש של $G$ כך ש- $\sum_{e\in T}w(e)>\sum_{e\in T'}w(e)$ ולכן,

$$-\sum_{e\in T}w'(e)=\sum_{e\in T}w(e)>\sum_{e\in T'}w(e)=-\sum_{e\in T'}w'(e)$$

מכאן, $\sum_{e\in T}w'(e)<\sum_{e\in T'}w'(e)$ בסתירה לכך ש- $T$ הוא עץ פורש מינימאלי של $G$ תחת פונקצית המשקל $w'$.

## שאלה 3:

יהי $G=(V,E)$ גרף לא מכוון, ממושקל וקשיר, ותהי $w:E\to R$ פונקצית המשקל על הצלעות.
בהינתן צלע $e\in E$, נגדיר את $MST_e$ של $G$ להיות עץ פורש של $G$ אשר מכיל את $e$ והוא מינימאלי במשקלו מכל העצים הפורשים של $G$ אשר מכילים את $e$.
תארו אלגוריתם, אשר בהינתן צלע $e\in E$ מחשב את $MST_e$.

## פתרון:

כשאנו באים לפתור בעיה שנראית מאוד דומה לבעיה אחרת (מציאת עץ פורש מינימאלי), אנו מיד חושבים על שני פתרונות אפשריים – רדוקציה לבעיה דומה או שינוי אלגוריתם ידוע כך שיתאים לבעיה שלנו. נראה את שתי האפשרויות לבעיה 1.

### פתרון ברדוקציה:

הפתרון ברדוקציה מאוד אינטואיטיבי. הצלע $e=(u,v)$ חייבת להיות בעץ פורש שאנו מחפשים, כך שאנו מחפשים למזער את עלות שאר הצלעות. הרעיון הוא "להעלים" את הצלע $e$ מהגרף ואז למצוא עץ פורש מינימאלי בגרף שנשאר. (רעיון אחר הוא לתת לצלע זו משקל קטן מכל שאר הצלעות ($-\infty$ למשל) ובזה לאלץ כל אלגוריתם שמחשב עץ פורש מינימאלי לבחור בצלע זו).

<u>איך "מעלימים" צלע?</u>
"נכווץ" את הקודקודים $u$ ו-$v$ לקודקוד חדש $x$. ז"א נוריד את הקודקודים $u$ ו-$v$, נוסיף קודקוד חדש $x$ ולכל קודקוד $y$ כך ש-$(u,y)\in E$ או $(v,y)\in E$ נוסיף צלע $(x,y)$. נקרא לגרף החדש $G'$.

<u>האם ניתן להפעיל את הקופסא השחורה?</u>
עוד לא! תיזכרו בהוכחת נכונות הרדוקציה.
מהו הקלט לכל אלגוריתם שמחשב עץ פורש מינימאלי? אלגוריתם כזה מצפה לקבל גרף פשוט וקשיר בקלט ובגרף שיצרנו, $G'$, יש כמה צלעות בין אותו זוג קודקודים.

יהי גרף $G''$, גרף פשוט שמתקבל מ-$G'$ ע"י כך שעבור כל זוג קודקודים עם מספר צלעות, נשאיר רק את הצלע עם המשקל המינימאלי.
עתה ברור כי $G''$ פשוט וקשיר. אמנם אין צורך להסביר, **חשוב מאוד לציין!**
נפעיל אלגוריתם כלשהו למציאת העץ הפורש המינימאלי של $G''$, $T''$.
מהו שלב תרגום הפלט? (שאלה לכיתה).

<u>מה נשאר להוכיח?</u>
<u>טענה ראשית</u>: הגרף המתקבל מ-$T''$ ע"י שחזור הקודקודים $u,v$ והצלע $e$ (וכמובן הורדת $x$) הוא עץ פורש של $G$, מכיל את $e$, ומינימאלי במשקלו מבין כל העצים הפורשים של $G$ אשר מכילים את הצלע $e$.

**<u>פתרון ע"י אלגוריתם דומה:</u>**

נשנה את אלגוריתם *Kruskal* למציאת עץ פורש מינימאלי כך שיתאים לבעיה שלנו.
אלגוריתם למציאת $MST_e$:

בהינתן גרף $G=(V,E)$ וצלע $e=(u,v)$

- אתחול: $E_T \leftarrow \{e\}$ , $E \leftarrow E \setminus \{e\}$
- כל עוד $E \neq \phi$
  - $e' \rightarrow$ הצלע בעלת משקל מינימאלי ב- $E$.
  - $E \rightarrow E \setminus \{e'\}$
  - אם $e'$ לא סוגרת מעגל ב- $T=(V,E_T)$
    - $E_T \leftarrow E_T \cup \{e'\}$.
- החזר את $T=(V,E_T)$.

<u>הוכחת נכונות האלגוריתם (</u>בהנחה ש- $G$ קשיר)

<u>טענה ראשית:</u>
כאשר תנאי סיום הלולאה מתקיים, אזי $T=(V,E_T)$ הוא $MST_e$ של $G$.

<u>טענה נשמרת (אינווריאנטה):</u>
תהי $E_i$ קבוצת הצלעות שנבחרה עד תחילת האיטרציה ה- $i$.
אזי לכל איטרציה $i$ קיים $MST_e$ המכיל את $E_i$.

<u>הוכחת הטענה הראשית על סמך הטענה הנשמרת:</u>
האלגוריתם עוצר אחרי |E| איטרציות.
לפי הטענה הנשמרת, ישנו $MST_e$ המכיל את $E_T$. כמו כן $T=(V,E_T)$ הוא עץ (כי הוא קשיר וחסר מעגלים) ולכן $T=(V,E_T)$ הוא $MST_e$.

<u>הוכחת הטענה הנשמרת:</u>
באינדוקציה על $i$.
**בסיס**: ברור שהטענה נכונה עבור $E_1=\{e\}$ כי כל $MST_e$ מכיל את $e$.

**הנחת האינדוקציה**:
נניח כי בתחילת האיטרציה ה- $i$ קיים $T=(V,E_T)$, $MST_e$ המכיל את $E_i$.

**צעד האינדוקציה:**

תהי $e'=(x,y)$ הצלע הנשקלת באיטרציה ה- $i$.

אם $e'$ סוגרת מעגל ב- $E_i$ אזי $E_{i+1}=E_i\subseteq E_T$ וסיימנו. (לא הוספנו צלע ונשארנו עם אותו העץ $E_T$).

אם $e'$ לא סוגרת מעגל ב- $E_i$ וגם $e'\in E_T$, אזי $E_{i+1}=E_i\cup\{e'\}\subseteq E_T$ וסיימנו. (זהו מקרה פשוט בו $E_T$ כבר מכיל את הצלע שבחרנו)

אחרת, $e'$ לא סוגרת מעגל ב- $E_i$ ו- $e'\notin E_T$ ולכן $E_{i+1}=E_i\cup\{e'\}$.

- כיוון ש- $e'$ לא סוגרת מעגל ב- $E_i$, לא קיים מסלול מ- $x$ ל- $y$ ב- $E_i$.
- כיוון ש- $T$ קשיר, חייב להיות מסלול מ- $x$ ל- $y$ ב- $T$. ולכן מסלול זה חייב להכיל צלע $e''=(a,b)$ שאינה ב- $E_i$ (זאת כי מהנחתנו לא קיים מסלול מ- $x$ ל- $y$ ב- $E_i$).
- נבחין כי $w(e')\le w(e'')$ כי גם $e''$ לא סוגרת מעגל ב- $E_i$ (אם הייתה סוגרת מעגל, הרי שיש מעגל ב- $T$, בסתירה לכך שזהו עץ), ולכן אם $w(e'')<w(e')$ האלגוריתם היה בוחן את $e''$ לפני $e'$ ומוסיף אותה ל- $E_j$ עבור $j\le i$.
- יהי $T'=(V,E'_T)$, כאשר $E'_T=E_T\setminus\{(a,b)\}\cup\{(x,y)\}$.

נוכיח כעת כי $T'$ הוא $MST_e$:

עפ"י משפט 2, $T'$ הוא עץ פורש.

כמו כן $w(T')=w(T)-w(e'')+w(e')\le w(T)$ כיוון ש- $w(e')\le w(e'')$ ולכן $T'$ הוא $MST_e$.

<u>שימו לב למבנה ההוכחה שראינו:</u>

1. המבנה מתאים לסכימת ההוכחה לאולגוריתם חמדן שראינו בתרגול קודם.
2. כל ההוכחה התבססה על טענה נשמרת (אינווריאנטה) שהוכחנו לאורך על ריצת האלגוריתם. את הטענה הראשית הוכחנו משילוב של טענה נשמרת ותנאי לולאה שלא התקיים. הרבה פעמים כדי להוכיח נכונות של לולאה אפשר:
   - נגדיר את האינווריאנטה.
   - נוודא <u>שהאינווריאנטה</u> מתקיימת לפני האיטרציה הראשונה.
   - נוכיח <u>שהאינווריאנטה</u> נשמרת מאיטרציה לאיטרציה, כל עוד תנאי הלולאה מתקיים.
   - נסיק את טענה הראשית מתוך האינווריאנטה ותנאי הלולאה שלא התקיים.

## <u>שאלה 3.2</u> (למחשבה)

תארו אלגוריתם מבוסס על *prim*, אשר בהינתן צלע $e\in E$ מחשב את $MST_e$.

## שאלה 4:

יהי גרף $G=(V,E)$ קשיר כך שכל משקלות קשתותיו שונים זה מזה. הוכח את הטענה הבאה: ל-$G$ קיים עץ פורש מינימאלי יחיד .

## פתרון:

<u>טענת עזר</u>: אם $T_1=(V,E_1),T_2=(V,E_2)$ הם שני עצים פורשים **שונים** לגרף $G=(V,E)$ אזי מתקיים $E_1\setminus E_2\neq\varnothing$ וגם $E_2\setminus E_1\neq\varnothing$ .

<u>הוכחת הטענה המרכזית</u>: נניח בשלילה שקיימים שני עצים פורשים מינימאליים שונים של $G=(V,E)$, $T_1=(V,E_1)$ ו- $T_2=(V,E_2)$.

נגדיר $H=(E_1\setminus E_2)\cup(E_2\setminus E_1)$ . לפי טענת העזר הקבוצה $H$ אינה ריקה. תהי $e\in H$ הצלע בעלת המשקל המינימאלי ב- $H$, ונניח ב.ה.כ כי $e\in E_1\setminus E_2$ .

נוסיף את $e$ ל- $T_2$, ונסמן גרף זה ב-$T_2'$. לפי משפט 2 ב $T_2$ קיים מעגל. $e$ בהכרח צלע במעגל כזה, כיוון שידוע כי ללא $e$ לא קיים מעגל ב- $T_2$. יהי $C$ מעגל ב-$T_2'$ המכיל את הצלע $e$. תהי $e'\in C$ כך ש- $e'\in E_2\setminus E_1$. $e'$ כזו קיימת, שכן אילולא הייתה קיימת אזי כל צלעות $C$ היו מוכלות ב-$E_1$, וזה אומר ש-$T_1$ מכיל מעגל, בסתירה להיותו עץ.

יהי $T'=(V,E')$ כך ש- $E'=(E_2\cup\{e\})\setminus\{e'\}$ . עפ"י משפט 2, $T'$ הוא עץ פורש של הגרף $G$ . מכיוון שגם $e$ וגם $e'$ שייכות ל- $H$, והקשת $e$ בעלת משקל מינימאלי מבין קשתות $H$, ניתן להסיק כי $w(e)\le w(e')$, ומכיוון שמנתון השאלה $w(e)\neq w(e')$, נסיק כי $w(e)<w(e')$ . לכן, $w(T')=w(T_2)+w(e)-w(e')\underset{w(e)-w(e')<0}{\Rightarrow} w(T')<w(T_2)$ .

לכן, $T'$ עץ פורש ל- $G$ בעל משקל קטן יותר משל העץ הפורש המינימאלי $T_2$ - סתירה.

<u>הוכחת טענת העזר</u>: $T_1=(V,E_1),T_2=(V,E_2)$ הם שני העצים הפורשים שונים לגרף $G=(V,E)$ לכן וודאי אחד התנאים הבאים מתקיים $E_1\setminus E_2\neq\varnothing$ או $E_2\setminus E_1\neq\varnothing$ . נניח בלי הגבלת הכלליות ש- $E_1\setminus E_2\neq\varnothing$ מכיוון ש- $|E_1|=|E_2|=|V|-1$ (ממשפט 1) לא יתכן ש-$E_2\subset E_1$ ולכן גם $E_2\setminus E_1\neq\varnothing$ מתקיים.

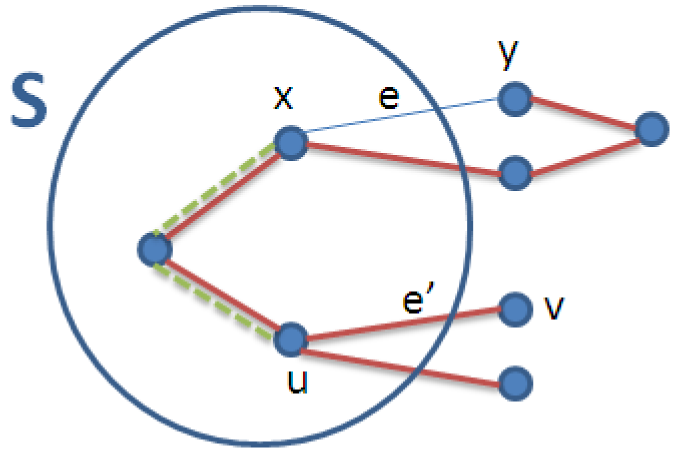
## **שאלה 5:**

להלן מוצע טיעון אחר להוכחת אלגוריתם PRIM. טיעון זה הינו **שגוי**.
עליכם להמחיש זאת ולאתר את השגיאה על ידי תיאור גרף שעליו הטיעון נכשל. יש להסביר מדוע הטיעון נכשל בגרף שתיארתם.

בהוכחת PRIM שהוצגה בכתה נתמקד במקרה המרכזי. כלומר, ישנה קבוצת קודקודים $S$ שנפרשו עד עתה על ידי החמדן, קבוצת צלעות $A$ שנבחרה עד כה על ידי החמדן, צלע $e=(x,y)$ הנבחרת באיטראציה הנוכחית, ועץ פורש מינימום $T=(V,F)$ כך ש $A\subseteq F$, כאשר $T$ מושג מהנחת האינדוקציה. אנו במקרה שבו $e\notin F$. בטיעון שהוצג בכיתה, במקרה זה הראנו כיצד לבנות עץ פורש מינימום $T'=(V,F')$ כך ש $A\cup\{e\}\subseteq F'$. נציע בנייה אחרת עבור $T'$:

נשים לב כי עבור $e$ מתקיים כי $x\in S$ ו $y\in V\setminus S$. בנוסף, היות ו-$T$ קשיר קיימת צלע $e'=(u,v)\in F$ כך ש-$u\in S$ ו-$v\in V\setminus S$. הסרת הצלע $e'$ מ $T$ מגדירה שני תתי עצים. אם כך, הצלע $e$ מחברת שני תתי עצים אלו ולכן נובע כי $T'=(V,(F\cup\{e\})\setminus\{e'\})$ הינו עץ פורש של הגרף. בפרט מתקיים כי משקל של $T'$ הינו לכל היותר המשקל של $T$.

**<u>הסבר:</u>**
האלגוריתם נכשל כיוון שהוא לא בחר בצלע $e'=(u,v)$ עם $u\in S$ ו-$v\in V\setminus S$ כך <u>שהיא נמצאת במסלול מ-$x$ ל-$y$</u>. אם היינו מוודאים כי $e'$ נמצאת במסלול מ-$x$ ל-$y$, אז היא הייתה שייכת למעגל שנוצר כתוצאה מהוספת $e$ והסרתה הייתה בהחלט משיגה את המטרה. אחרת יכול להיווצר מצב בו הסרת $e'$ אכן מגדירה שני תתי עצים, אך $e$ היא לא הצלע שמחברת ביניהם.
דוגמא כזו הינה:




כאשר הצלעות המקווקוות הן הצלעות בקבוצה $A$, ואילו הצלעות המודגשות הן הצלעות של העץ הפורש $T$ הקיים מהנחת האינדוקציה.