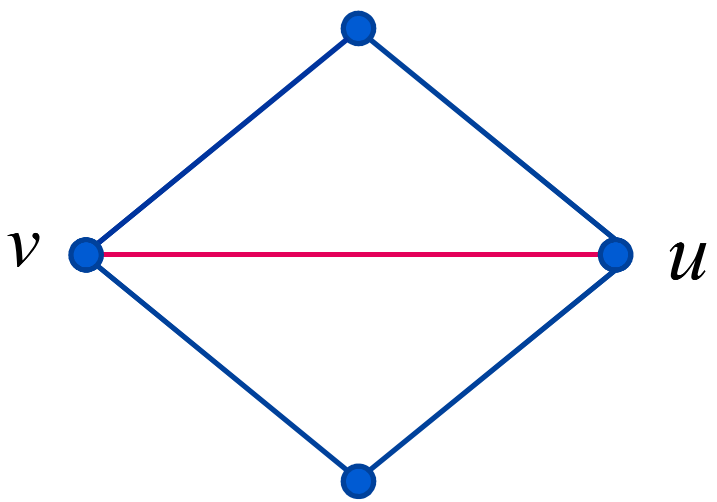
תרגול 4 – עפיימים

תכונה כללית של עצים

- ***טענה:*** אם נוסיף קשת לעץ פורש, היא תסגור <u>מעגל</u> יחיד.
- ***הוכחה:***
  - נסמן את הקשת שהוספנו בתור (v,u). היות שבעץ כבר היה מסלול בין v לבין u, הקשת בהכרח תסגור מעגל.
  - אם הקשת סוגרת שני מעגלים, נוכל ליצור מעגל מקשתות שני המעגלים, ללא (v,u). בסתירה לכך שמדובר בעץ.

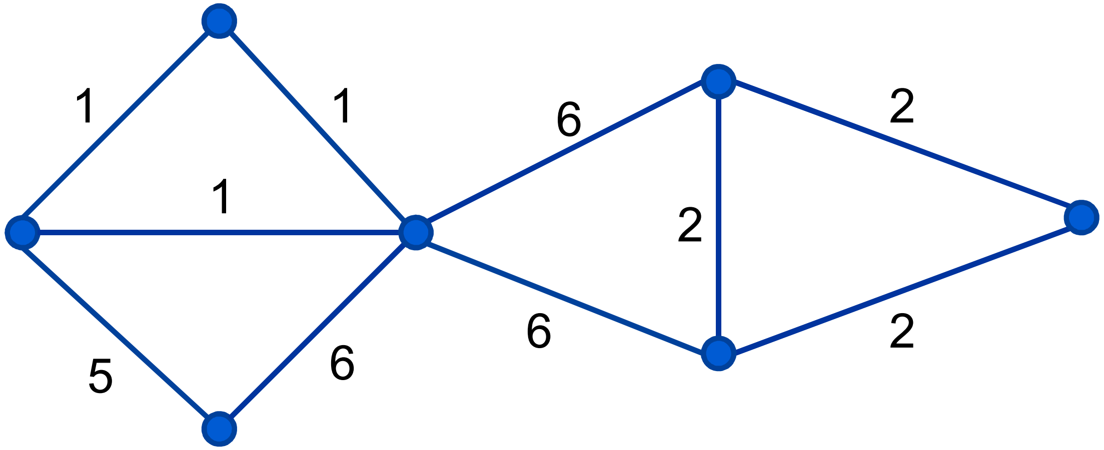


שאלה 1 – תכונות של עפ״מים

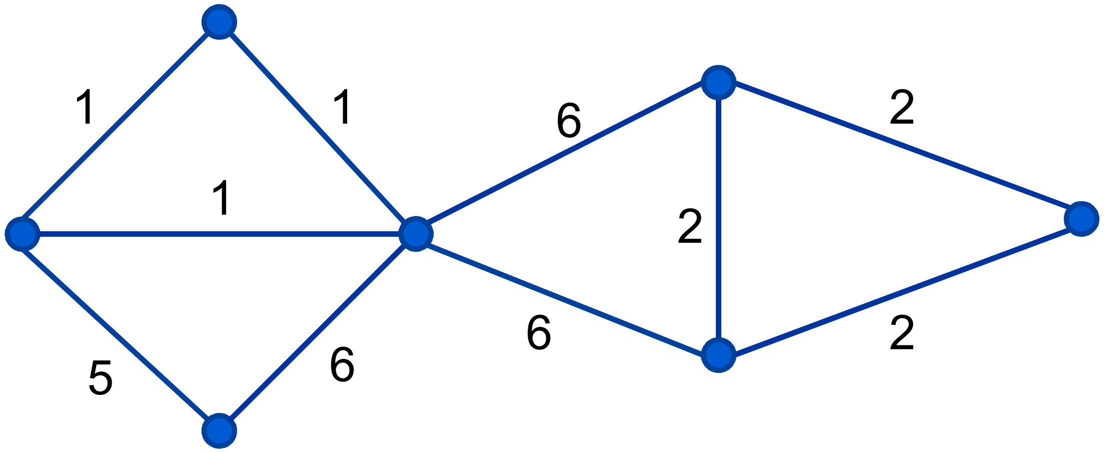
- **תרגיל:** נתונים גרף לא מכוון וקשיר $G=(V,E)$, פונקצית משקל $w:E\to R$, ושני עפ"מים T_1,T_2.
הוכיחו שהטענה הבאה מתקיימת עבור כל מספר ממשי:
תהי $E_{1,r}$ קבוצת הקשתות של T_1 שמשקלן לכל היותר r.
הקבוצה $E_{2,r}$ מוגדרת באופן סימטרי. הגרפים $(V,E_{1,r})$ ו- $(V,E_{2,r})$ מחזיקים בדיוק באותם רכיבי קשירות.

שאלה 1 – דוגמא

T_1



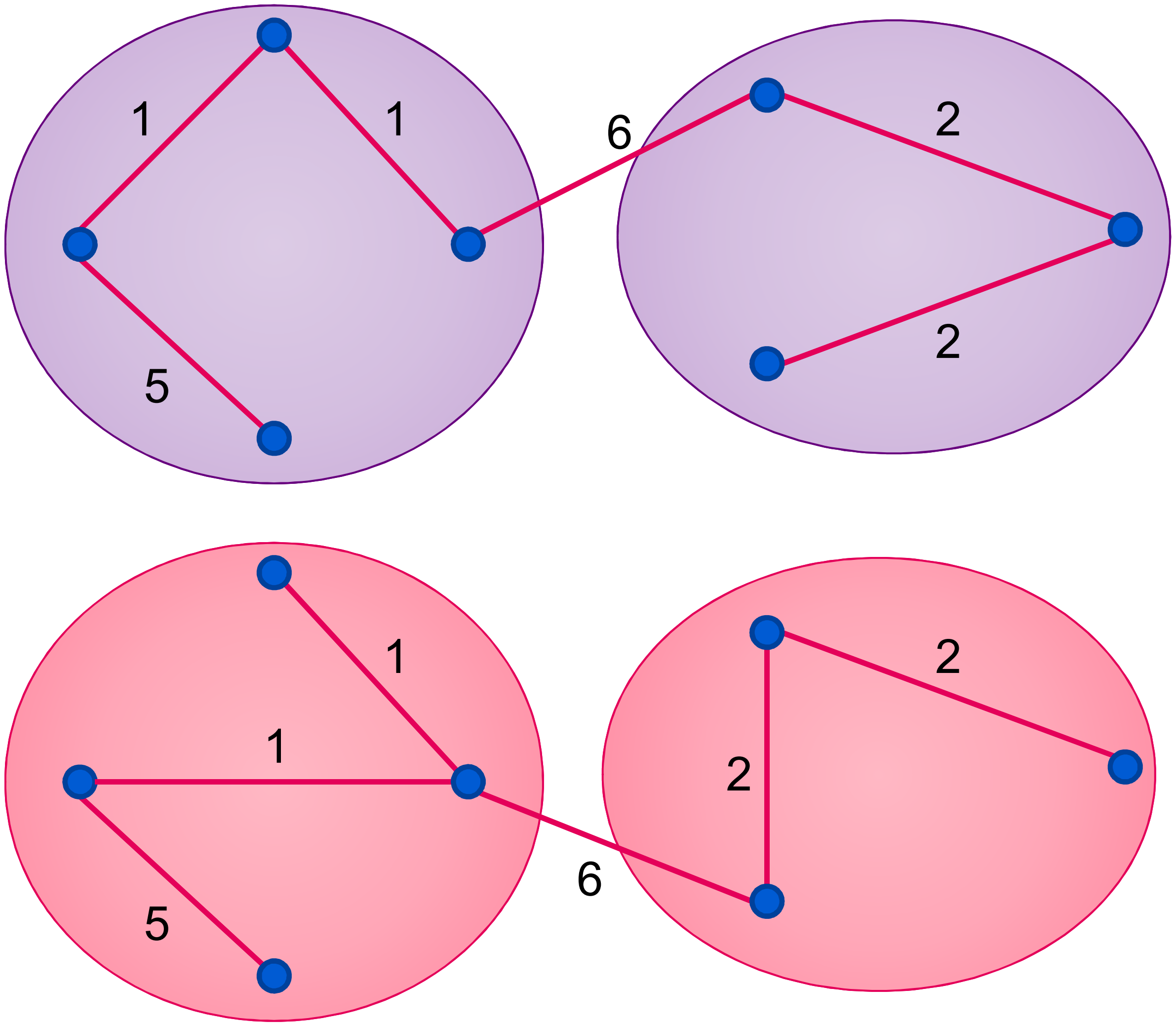
T_2



שאלה 1 – דוגמא

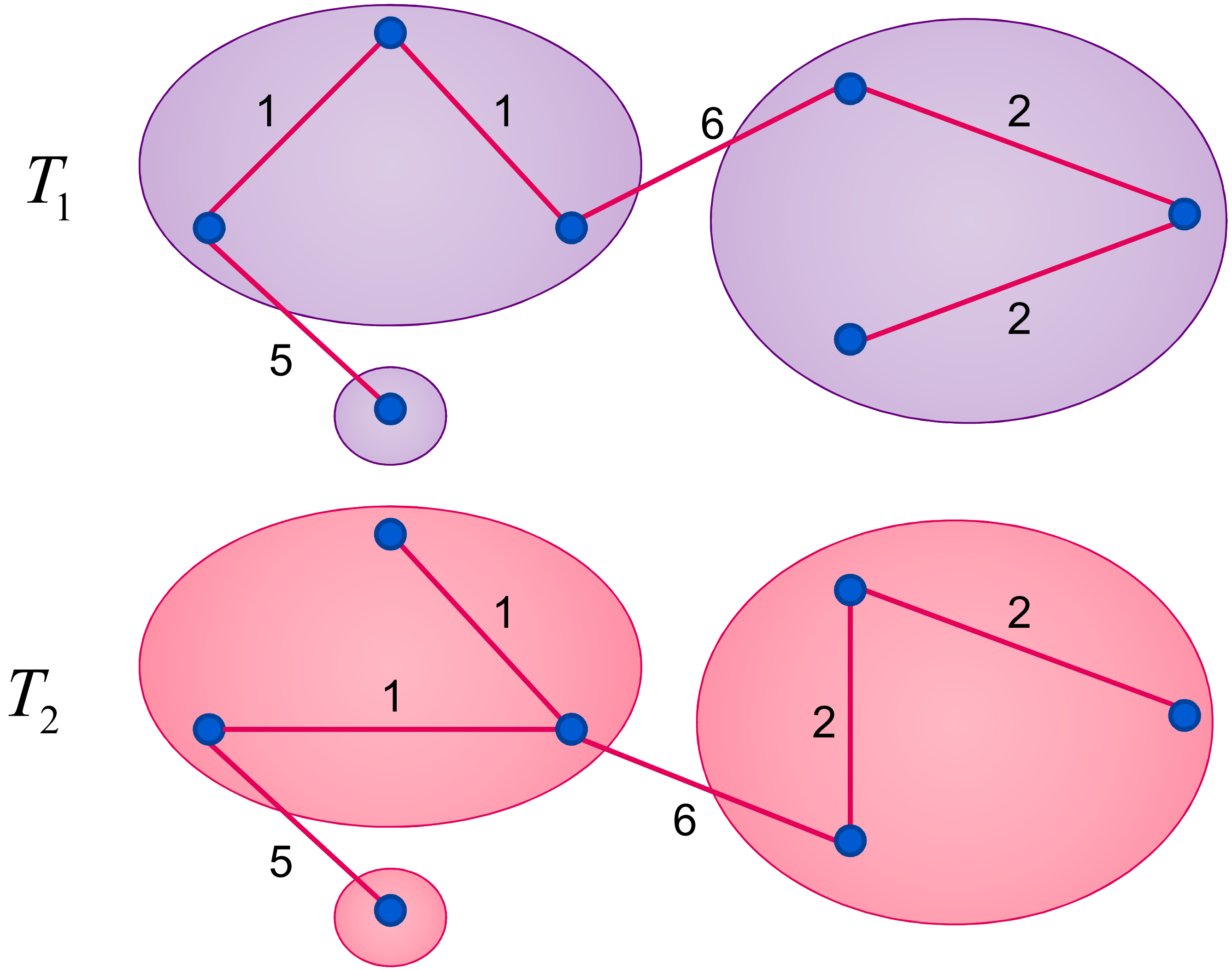
T_1

T_2



$r = 5$

שאלה 1 – דוגמא



$r = 3$

שאלה 1 – פתרון

- נניח בשלילה שקיימים r ושני עפ"מים T_1, T_2, כך שרכיבי הקשירות של $G_1 = \left(V, E_{1,r}\right)$ ושל $G_2 = \left(V, E_{2,r}\right)$ אינם זהים:

קיימים שני קודקודים הנמצאים באותו רכיב קשירות בגרף אחד, וברכיבים שונים בגרף השני.

בה"כ, נניח שקיימת קשת $e = \left(v, u\right)$ ב- G_1, כך ששני הקודקודים שלה נמצאים ברכיבי קשירות שונים ב- G_2. (הוכיחו!)

אם נוסיף את e לעץ T_2, היא תסגור מעגל שלפחות אחת מקשתותיו במשקל גדול מ- r. נסמן קשת כזו בתור e'.

כלומר, אם נחליף את e' על ידי e, נקבל עץ קל יותר. בסתירה לכך ש- T_2 הינו עפ"מ.

למה המעגל בהכרח מכיל קשת כבדה?

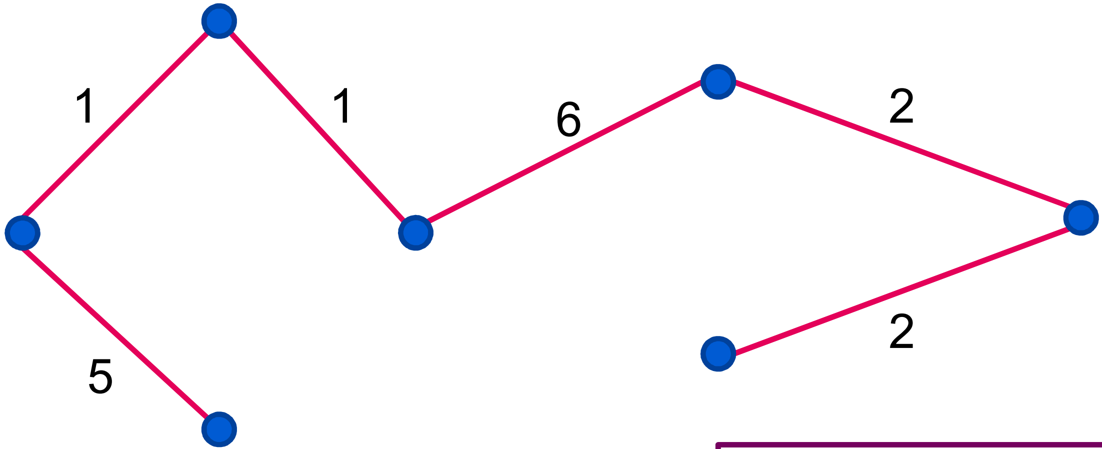
- ב- G_1 קיימת הקשת (v,u), ובב- G_2 הקודקודים v ו- u נמצאים ברכיבי קשירות שונים.
- ב- T_2 יש מסלול p בין v ו- u. כאשר אנו מוסיפים את הקשת (v,u) היא סוגרת מעגל יחד עם p.
- היות ש- v ו- u נמצאים ברכיבי קשירות שונים בגרף G_2, לפחות אחת מהקשתות של p הינה במשקל גדול מ- r.

שאלה 2 - רשימות משקלים

- **תרגיל:** נתונים גרף לא מכוון וקשיר $G=(V,E)$, פונקציית משקל $w:E\to R$, ושני עפ"מים T_1,T_2. תהי $\alpha_1\le\alpha_2\le\ldots\le\alpha_{n-1}$ רשימה ממוינת של משקלי קשתות T_1. באופן דומה, תהי $\beta_1\le\beta_2\le\ldots\le\beta_{n-1}$ רשימה ממוינת של משקלי קשתות T_2. הוכיחו שמתקיים $\forall i:\alpha_i=\beta_i$.

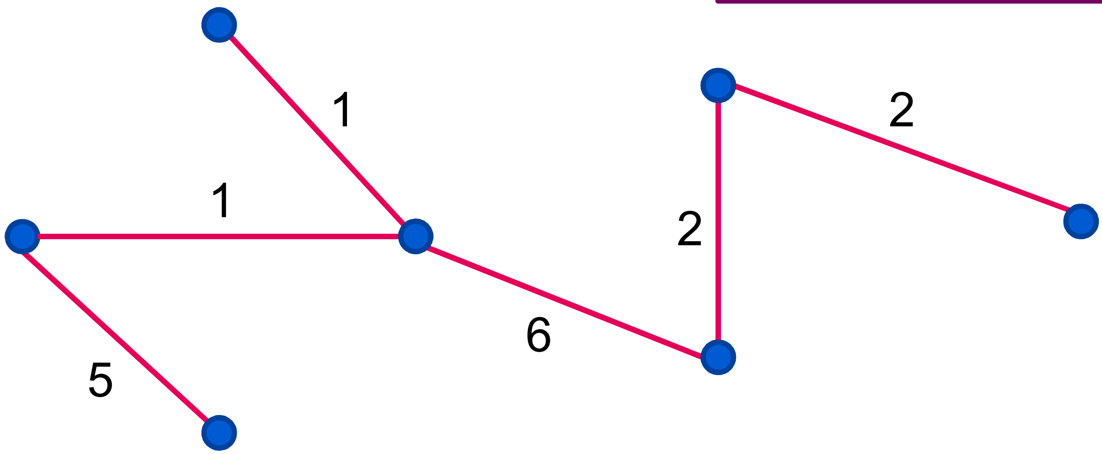
שאלה 2 – דוגמא (על הגרף משאלה 1)

T_1



משקלי הקשתות : 1,1,2,2,5,6

T_2



שאלה 2 – פתרון

- תהי $W=\{w_1,w_2,\ldots\}$ רשימה ממוינת של המשקלים בגרף. נוכיח את הטענה באינדוקציה עבור המשקלים שאינם גדולים מ- w_i:

  בסיס האינדוקציה – נראה שבשני העצים יש אותו מספר קשתות במשקל w_1.

  מהשאלה הקודמת עם $r=w_1$, עבור שני העצים, הקשתות במשקל זה יוצרות את אותם רכיבי קשירות. לכן, מספר הקשתות זהה.

שאלה 2 – פתרון

- צעד האינדוקציה – נניח שרשימות המשקלים זהות עד האיברים בגודל w_{i-1}, ונראה שיש בשני העצים אותו מספר של קשתות במשקל w_i:

לפי השאלה הקודמת עם $r = w_i$, אם נזרוק את הקשתות שגדולות ממשקל זה, נקבל בשני העצים את אותם רכיבי הקשירות.

זאת אומרת שיש בשני העצים אותו מספר קשתות במשקל $\geq w_i$.

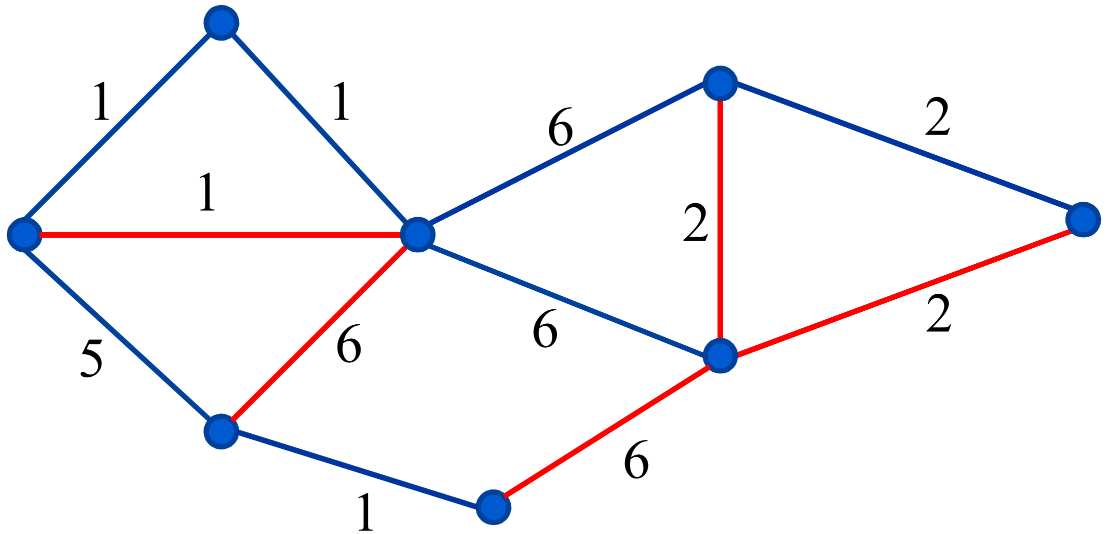
(כי בעץ עם k קודקודים יש $k-1$ קשתות.)

מהנחת האינדוקציה, יש בשני העצים אותו מספר קשתות עם משקל $\geq w_{i-1}$. לכן, יש בשני העצים אותו מספר של קשתות במשקל w_i.

שאלה 3 – קשתות צבועות

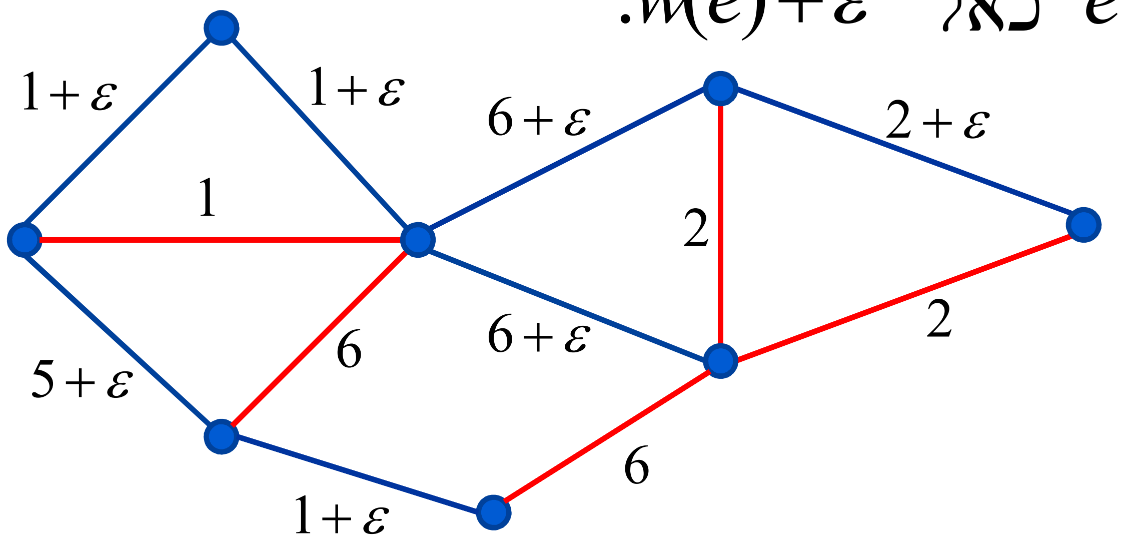
- **תרגיל:** נתונים גרף לא מכוון וקשיר $G=(V,E)$, ופונקציית משקל $w:E\rightarrow R$. בנוסף, כל קשת צבועה בכחול או באדום. תארו אלג' שמוצא עפ"מ המכיל מספר מקסימאלי של קשתות אדומות (מבין העפ"מים).

שאלה 3 – דוגמא



רעיון לפתרון

- נרצה להעדיף קשתות אדומות.
- נגדיר קבוע קטן ε ונתייחס למשקל של קשת כחולה e כאל $w(e)+\varepsilon$.

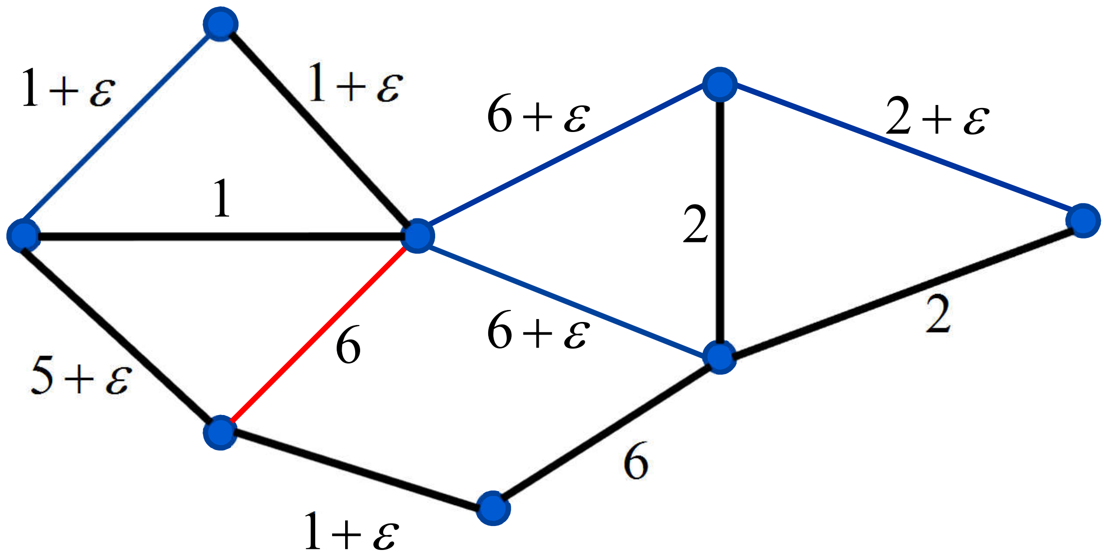


- כלומר, $w': E \rightarrow R$ היא פונקצית משקל חדשה, כך שלכל קשת אדומה, $w'(e)=w(e)$, ולכל קשת כחולה $w'(e)=w(e)+\varepsilon$.

נכונות הרעיון

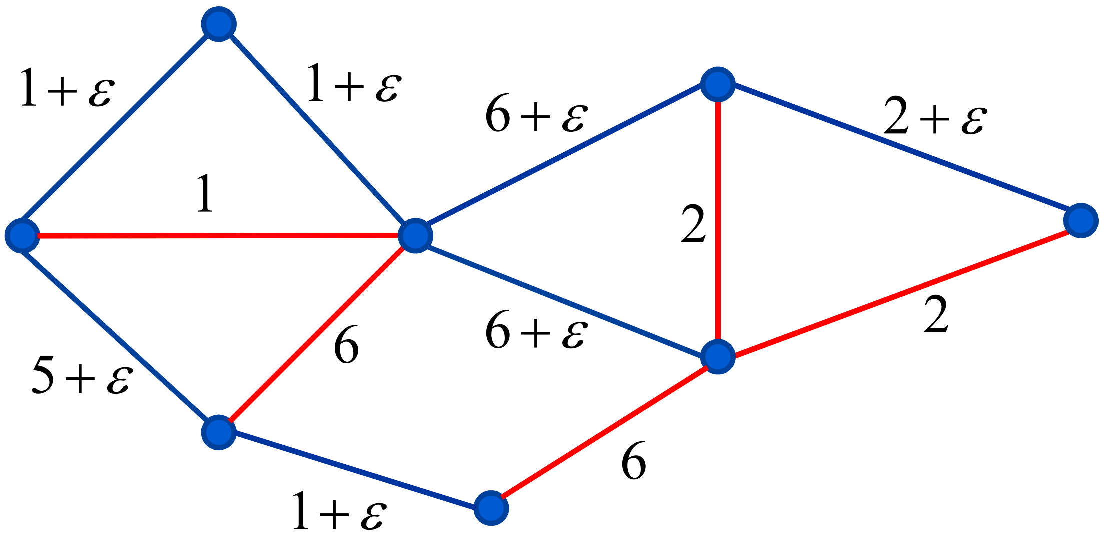
- נניח ש ε **מאוד** קטן. (נסביר עוד מעט בדיוק כמה קטן).
- נריץ את האלגוריתם של Kruskal על הגרף עם פונקצית המשקל החדשה, ונמצא עפ"מ.

$1 \quad 1+\varepsilon \quad 1+\varepsilon \quad 1+\varepsilon \quad 2 \quad 2 \quad 2+\varepsilon \quad 5+\varepsilon \quad 6 \quad 6 \quad 6+\varepsilon \quad 6+\varepsilon$



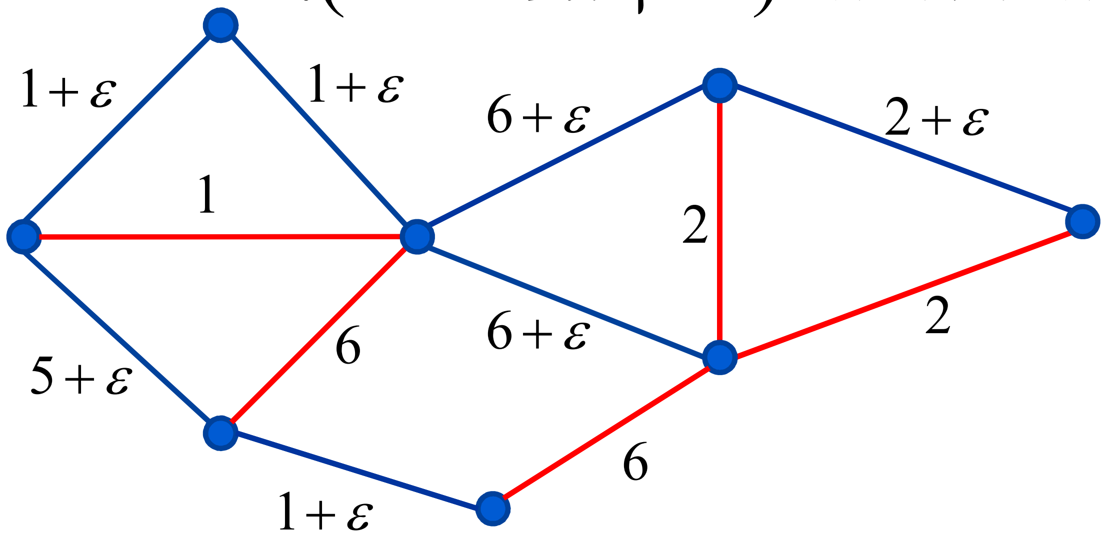
פורמלית

- נגדיר פונקצית משקל חדשה $w': E \rightarrow R$, כאשר לכל קשת אדומה $w'(e) = w(e)$, ולכל קשת כחולה $w'(e) = w(e) + \varepsilon$.



הוכחת הטענה

- **טענה:** עפ"מ ב G עם w' הוא גם עפ"מ ב G עם w.
- **הוכחה:** שיעורי בית.
- מתרגיל 2, כיוון שהאלגוריתם של Kruskal מוצא עפ"מ ב - G עם w', הוא מוצא עפ"מ ב G המכיל מספר מקסימאלי של קשתות אדומות (מבין העפ"מים).

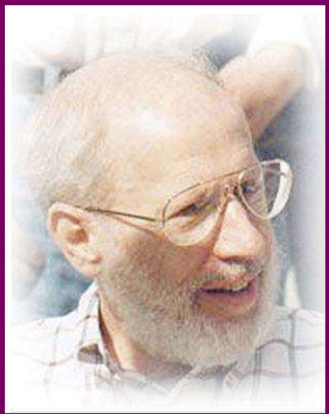


עד כמה קטן $\boxed{\varepsilon}$ צריך להיות?

- אפשר לקחת $\varepsilon < \Delta_{min}$, כאשר Δ_{min} ההפרש הקטן ביותר בין משקלי 2 קשתות בגרף (שהוא לא 0).
  (שיעורי בית : לוודא שהטענה בעמוד הקודם מתקיימת עם הגדרה זו של ε).

- כמה זמן יקח לחשב את ε?

שאלה קטנה



- נניח שנחליף את האלג' של Kruskal באלג'
- של Prim, האם נכונות האלגוריתם תשמר?

כן

- האם זמן הריצה של האלג' ישמר?

בערך...

$$O(|E|+|V|\log|V|)$$

VS.

$$O(|E|\log|E|)$$

•

פתרון נוסף

- **פתרון:** במקום להוסיף ε לקשתות כחולות, נסדר את הקשתות לפי משקל ואז לפי צבע בהרצת קרוסקל.
- **תיאור ההוכחה:** האלגוריתם של קרוסקל בוחר את הקשתות לפי הסדר שהוא מקבל אותם. אם נבטיח שיקבל את הקשתות האדומות קודם, הוא תמיד יוסיף אותן אם יוכל.

- **מסקנה:** אם נרצה שעפ"מ מסויים יבחר, תמיד נוכל לסדר את הקשתות כך שקרוסקל יבחר אותו!

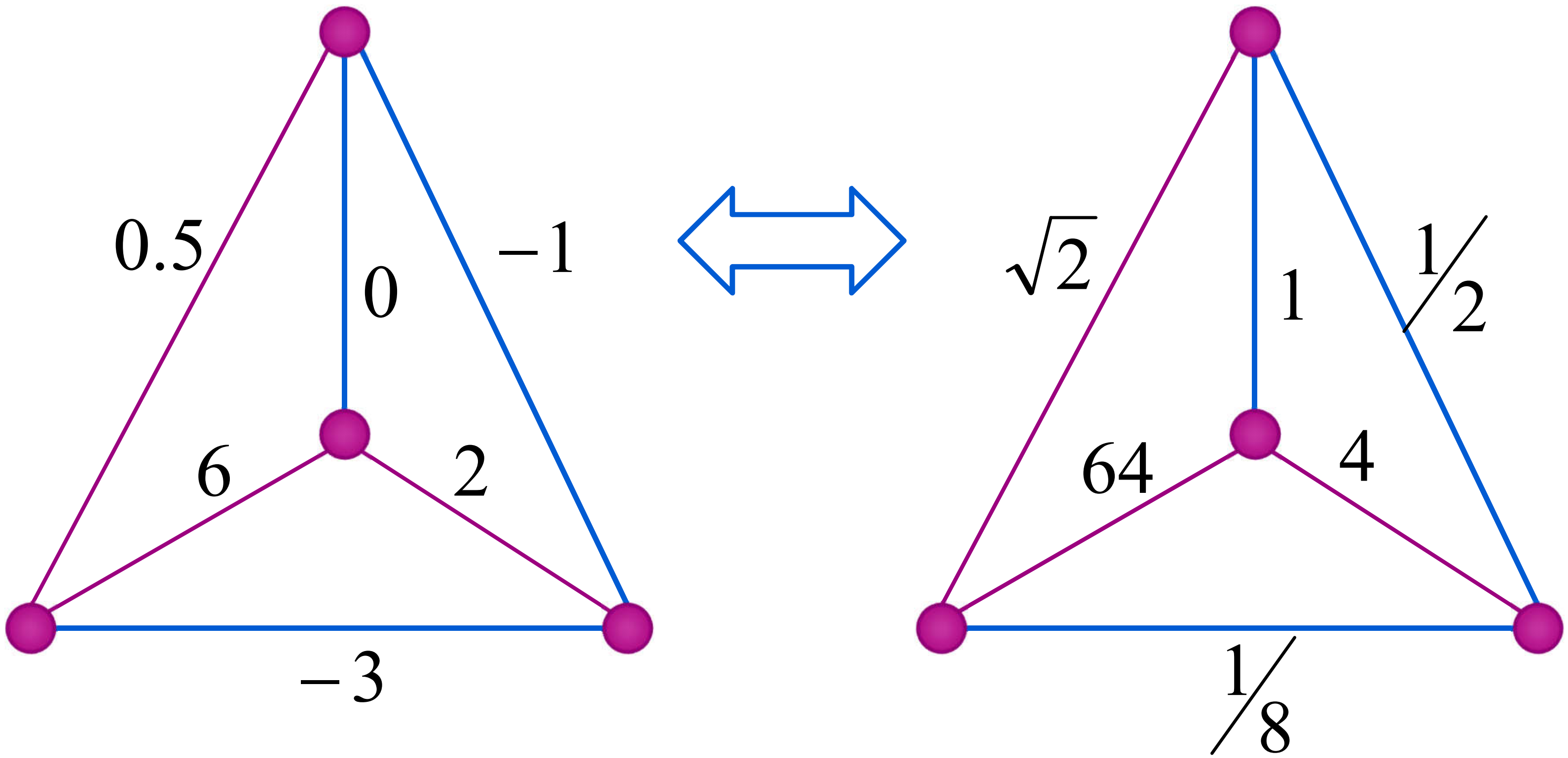
טענה

- יהי $G=(V,E)$ גרף קשיר ולא מכוון, תהי $w:E\rightarrow R$ פונקציית משקל, ו- $f:R\rightarrow R$ פונקציה מונוטונית עולה (ממש). תהי $w':E\rightarrow R$ פונקצית משקל כך ש $w'(e)=f(w(e))$.

טענה: T הוא עפ"מ של G ביחס ל- w אם"ם הוא עפ"מ ביחס ל- w'.

דוגמא

$$f(x) = 2^x$$

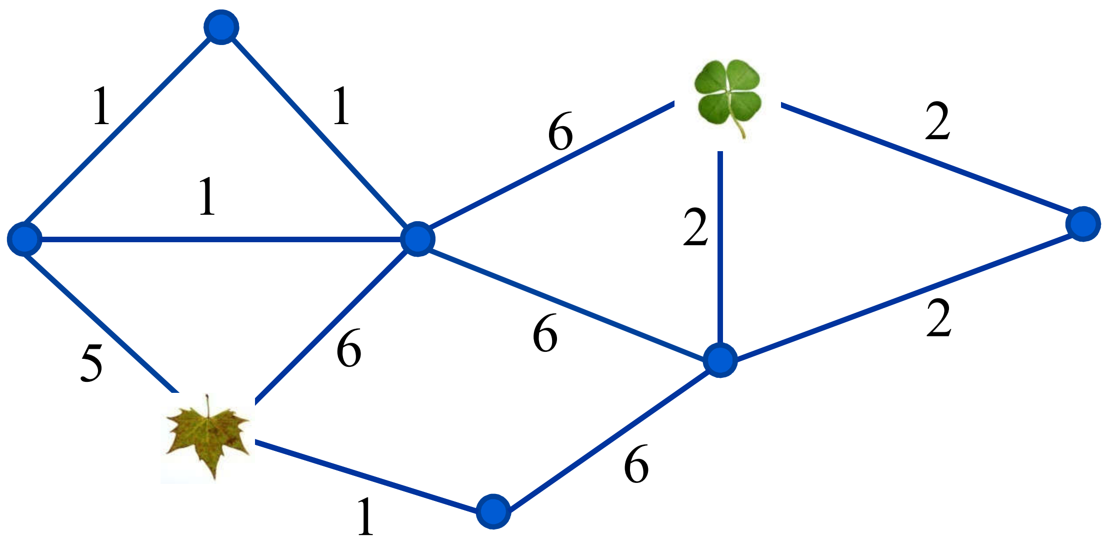


הוכחת הטענה

- נניח ש T עפ"מ של עץ ביחס ל - w . נסדר את הקשתות כך שקרוסקל ייצר את T . נפעיל את f על המשקולות ונקבל את אותו הסדר, ולכן הפעלת קרוסקל תייצר גם כן את T .
- ההוכחה בכיוון השני זהה.

שאלה 4 – עלים מנדטוריים

- **תרגיל:** נתונים גרף קשיר ולא מכוון $G=(V,E)$, פונקציית משקל $w:E\rightarrow R$, ותת קבוצה של קודקודים $L\subset V$. תארו אלג' המוצא עץ במשקל מינימאלי מבין העצים הפורשים שכל איברי L הם עלים שלהם (או מודיע שלא קיים עץ כזה).



שאלה 4 – פתרון

- נגדיר גרף $G'=\left(V',E'\right)$, כך ש- $V'=V\setminus L$ ו-E' הינה קבוצת הקשתות ששני הקודקודים שלהם ב- V'.
- האלגוריתם:

נבנה את G'.

אם הגרף אינו קשיר, נודיע שלא קיים עץ מתאים.

אחרת, נמצא עפ"מ T' ב- G'.

עבור כל קודקוד ב- L, נמצא את הקשת הקלה ביותר שמחברת בינו לבין T', ונוסיף אותה לעץ (אם לא קיימת קשת ל- T', נודיע שלא קיים עץ מתאים).

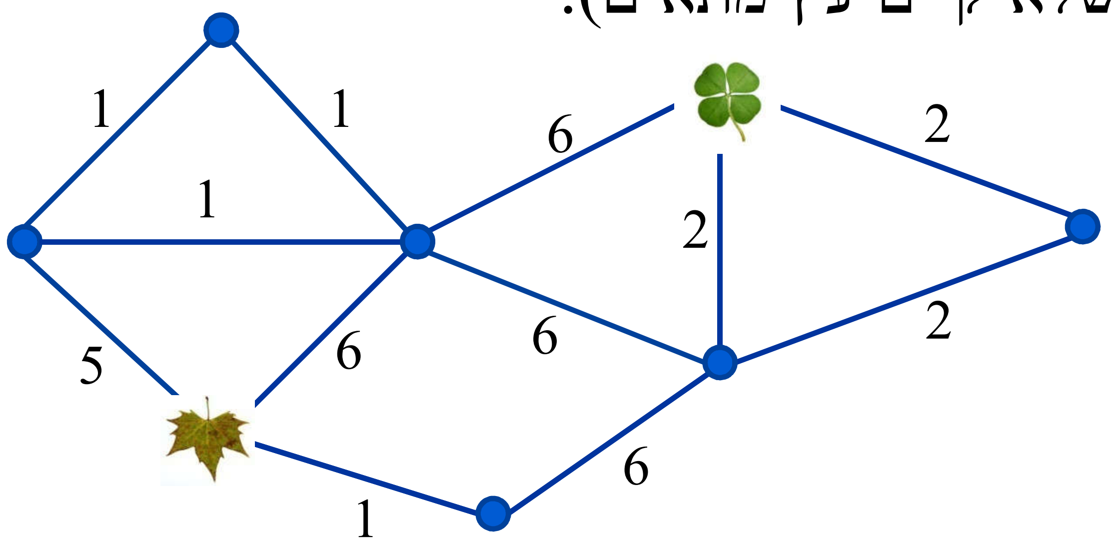
הרצה לדוגמא

נבנה את G'.

אם הגרף אינו קשיר, נודיע שלא קיים עץ מתאים.

אחרת, נמצא עפ"מ T' ב- G'.

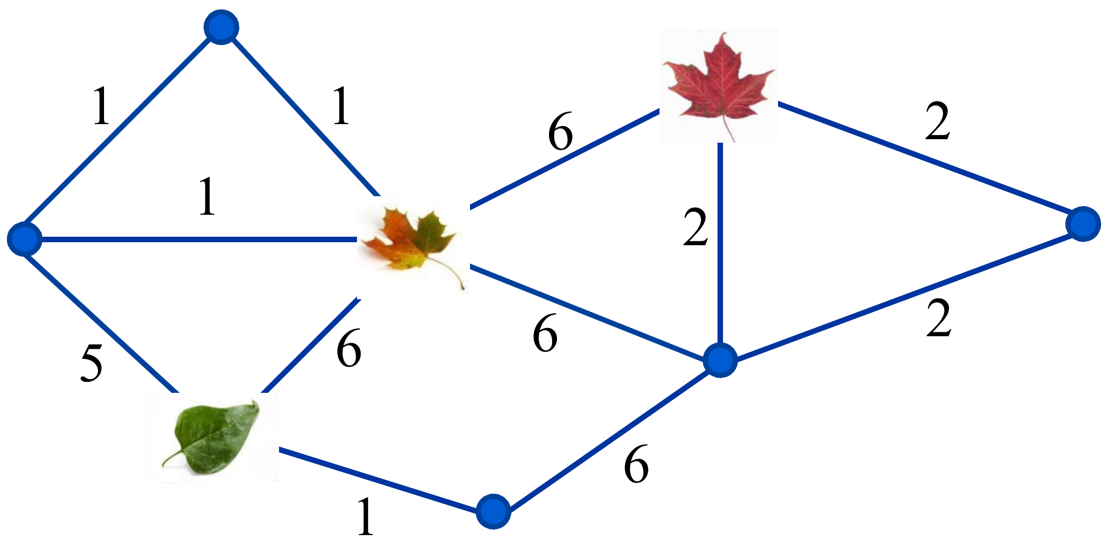
עבור כל קודקוד ב-L, נמצא את הקשת הקלה ביותר שמחברת בינו לבין T', ונוסיף אותה לעץ (אם לא קיימת קשת ל- T', נודיע שלא קיים עץ מתאים).



דוגמא נוספת

נבנה את G'.

אם הגרף אינו קשיר, נודיע שלא קיים עץ מתאים.



סיבוכיות

- נבנה את G'.

$$O(|V|+|E|)$$

- נמצא עפ"מ ב- G'. אם לא קיים עפ"מ כיוון שהגרף אינו קשיר, האלג' יודיע שלא קיים עפ"מ כמבוקש.

$$O(|E|\log|E|) \;\; / \;\; O(|E|+|V|\log|V|)$$

- עבור כל קודקוד ב- L:

נמצא את הקשת הקלה ביותר שמחברת בין הקודקוד לבין העץ שמצאנו בשלב השני, ונוסיף אותה לעץ.

$$O(|E|)$$

הוכחת נכונות

- נגדיר את התנאי C להיות – כל עלה מחובר בקשת ל- G'.
- נראה שקיים עץ פורש מתאים ב- G אם"ם מתקיימים שני התנאים שהאלג' בודק:

G' קשיר וגם מתקיים C $\Leftarrow$ קיים עץ פורש ב- G שאיברי L עלים בו.

קיים עץ פורש ב- G שאיברי L עלים שלו $\Leftarrow$ G' קשיר (אם נזרוק את העלים מהעץ, הגרף עדיין יהיה קשיר), וגם מתקיים C.

מינימאליות העץ הפורש

- נבחן את חיבור העלים לעץ:

מההגדרה, קודקוד $v \in L$ מחובר בעץ לקשת בודדת e.

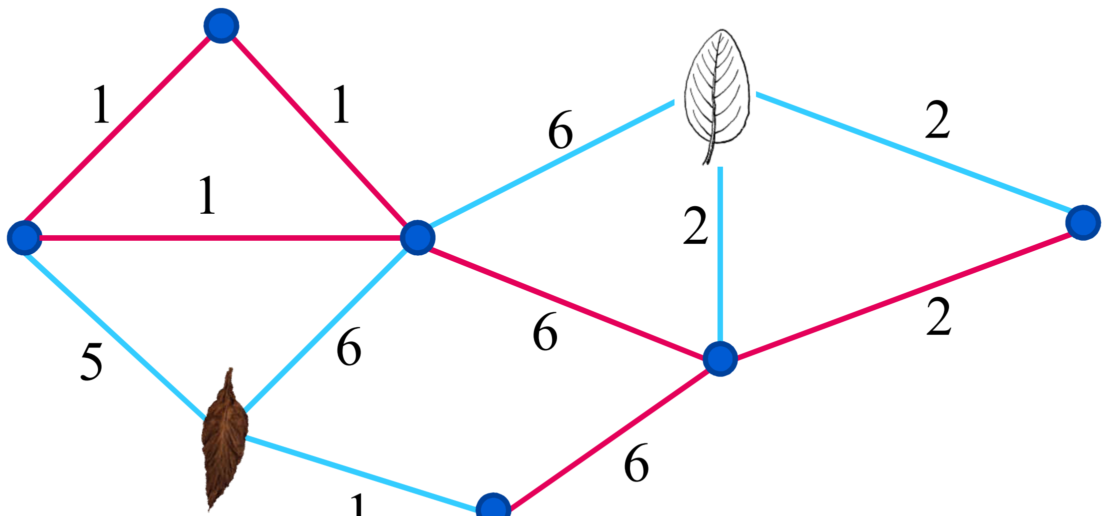
בנוסף, e מחוברת לקודקוד מ- V'.

בעץ המינימאלי, קודקוד $v \in L$ יהיה מחובר לשאר העץ דרך הקשת הקלה ביותר שמחברת אותו ל- V'.

אם זה אינו המצב, נוכל לקבל עץ קל יותר ע"י חיבור v לקשת קלה יותר.

מינימאליות העץ הפורש

- היות ואנו יודעים בדיוק איזה קשתות סמוכות לאיברי L,
נותר רק למצוא עפ"מ של G'.



שימוש מעניין בעפ״מים

Use of the Minimum Spanning Tree Model for Molecular Epidemiological Investigation of a Nosocomial Outbreak of Hepatitis C Virus Infection (Spada et al., 2003)

The MST model was applied to the molecular data and, together with the available clinical-epidemiological data, was used to identify the root of transmission of the outbreak and the most probable patient-to-patient chain of transmission. The basic concept of the model is that the outbreak and the most probable patient-to-patient chain of transmission are considered as a special graph connecting all patients with the minimum viral genetic distances among them.