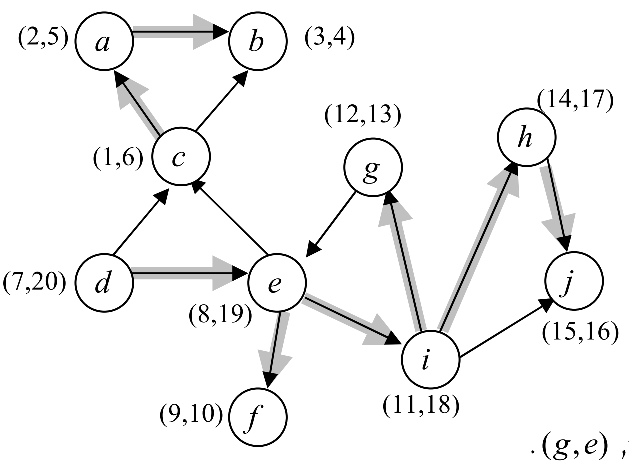
הטכניון – מכון טכנולוגי לישראל הפקולטה למדעי המחשב

אלגוריתמים 1 (234247) – סמסטר אביב 2005

תרגול 2: DFS

<u>דוגמת הרצה ל-DFS:</u>

להלן דוגמת הרצה של אלגוריתם DFS על גרף מכוון G (ליד כל צומת v מצויינים זמן הגילוי - $d[v]$ והנסיגה - $f[v]$).



תוצאת ההרצה היא יער DFS מכוון (הקשתות האפורות).

את קשתות G ניתן לסווג לשני סוגים:

- קשתות עץ. לדוגמה, (d,e).
- קשתות אחוריות (מצאצא לאב קדמון). לדוגמה, (g,e).
- קשתות קדמיות (מאב קדמון לצאצא). לדוגמה, (i,j).
- קשתות חוצות. לדוגמה, (e,c).

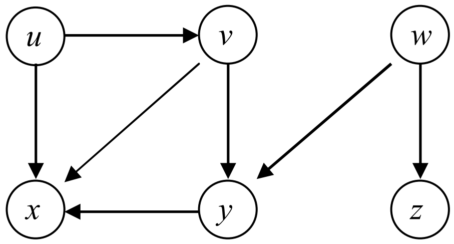
במקרה הלא-מכוון ישנן רק קשתות עץ וקשתות אחוריות.

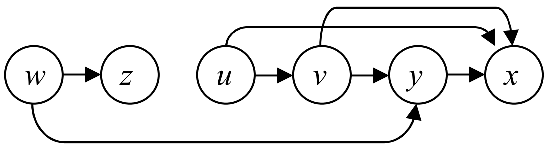
<u>תזכורת:</u>

- בעץ מכוון עם שורש r, כל צומת על המסלול היחיד מ-r לצומת v כלשהו נקרא **<u>אב קדמון</u>** של v. אם (w,v) היא הקשת האחרונה במסלול הנ"ל, אזי w נקרא **<u>אב</u>** של v.
- אם u הוא אב קדמון של v אזי נאמר כי v הוא **<u>צאצא</u>** של u. אם w הוא אב של v אזי נאמר כי v הוא <u>בן</u> של w.

משפט (מכונה "משפט המסלול הלבן" ב-CLRS): ביער ה-DFS של גרף (מכוון או לא-מכוון), צומת v הוא צאצא של צומת u אם ורק אם בזמן גילוי צומת u קיים מסלול מ-u ל-v דרך צמתים שעדיין לא התגלו (צמתים "לבנים").

מיון טופולוגי[1]

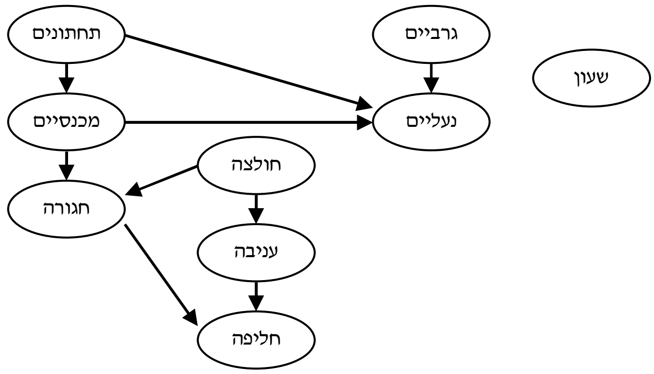




<u>*Directed Acyclic Graph – DAG – גרף מכוון חסר מעגלים.*</u>

<u>מיון טופולוגי</u> של DAG $G=(V,E)$ הוא סידור לינארי של קבוצת הצמתים V כך שלכל קשת $(u,v)\in E$ מתקיים ש-u מופיע לפני v בסדרה.
ניתן לראות מיון טופולוגי כסידור של הצמתים לאורך קו ישר כך שכל הקשתות המכוונות יהיו בכיוון זהה.
יתכן כי ל-DAG יהיו מספר סידורים כאלה, אך אם הגרף אינו DAG לא קיים סידור חוקי.

ניתן להשתמש ב-DAG לייצוג קדימויות בין אירועים שונים. הגרף בדוגמא הבאה (לקוח מתוך CLRS) מתאר את סדר הלבוש של פרופסור בבוקר. ישנם פריטי לבוש אותם יש ללבוש בסדר מסוים (לדוגמא, גרביים לפני נעליים) ואחרים שאינם תלויים זה בזה (לדוגמא, שעון ועניבה). מיון טופולוגי של גרף זה עשוי להציע סדר התלבשות מתאים.



בקורס מבני נתונים 1 נלמד אלגוריתם למיון טופולוגי אשר קובע את אחד המקורות בגרף[2] כצומת הבא בסידור, מסיר אותו ואת הקשתות היוצאות ממנו מהגרף וממשיך באופן דומה. אנו נראה אלגוריתם למיון טופולוגי של גרף מכוון חסר-מעגלים $G=(V,E)$ המבוסס על DFS:

1. הרץ DFS על G לחישוב זמן הסיום $f[v]$ לכל צומת $v\in V$.
2. לכל צומת $v\in V$, בזמן הנסיגה מ-v (קביעת $f[v]$), הכנס את v לתחילת רשימה מקושרת L.
3. סדר הצמתים ב-L הוא המיון הטופולוגי (הצומת שנכנס אחרון ל-L הוא הראשון בסידור וכו').

[1] מבוסס על פרק 22.4 ב-CLRS.

[2] מקור הוא צומת בגרף מכוון אליו אין קשתות נכנסות.

סיבוכיות: סיבוכיות הזמן של האלגוריתם היא $O(V+E)$ [3] מאחר ובמהלך ה-DFS הוספנו פעולה המתבצעת לכל צומת פעם אחת בדיוק ולוקחת $O(1)$.

הוכחת נכונות: לצורך הוכחת הנכונות נעזר בלמה הבאה.

למה (22.11): גרף מכוון G הוא חסר-מעגלים אם ורק אם בכל הרצת DFS על G לא מתקבלות קשתות אחוריות.

הוכחת הלמה:

$\Leftarrow$ יהא G גרף חסר-מעגלים. נניח בשלילה שקיימת הרצת DFS בה התקבלה קשת אחורית $(u \to v)$. זו קשת אחורית, כלומר v הוא אב קדמון של u. לכן, קיים מסלול מכוון מ-v ל-u בעץ ה-DFS, ובפרט ב-G, ולכן הקשת $(u \to v)$ סוגרת מסלול זה למעגל מכוון בסתירה לכך ש-G חסר מעגלים.

$\Rightarrow$ נניח ש-G מכיל מעגל C ונראה שכל הרצת DFS על G מניבה קשת אחורית. נסמן ב-v את הצומת הראשון על C שמתגלה בהרצת DFS על G. יהא u הצומת הקודם ל-v על C. בזמן גילוי v קיים "מסלול לבן" מ-v ל-u דרך צמתי C, לכן u צאצא של v ולכן $(u \to v)$ קשת אחורית. □

נמשיך בהוכחת נכונות האלגוריתם. נסתכל על הרגע בו קשת $e=(u \to v) \in E$ נסרקת במהלך ה-DFS, ונראה כי תמיד מתקיים כי $f[u] > f[v]$. אם e קשת עץ, כלומר u האב של v, אזי ברור כי $f[u] > f[v]$ (DFS נסוג מצומת רק לאחר שביצע נסיגה מכל בניו). אם e קשת קדמית, כלומר u אב קדמון של v, שוב $f[u] > f[v]$ כי ניתן להפעיל את הטעון הקודם על כל זוג צמתים עוקבים במסלול המכוון מ-u ל-v. e אינה יכולה להיות קשת אחורית עפ"י הלמה הקודמת, לכן נותר לבדוק רק את המקרה ש-e קשת חוצה. במקרה זה DFS כבר נסוג מ-v, כלומר $f[v]$ כבר נקבע ושוב מתקיים כי $f[u] > f[v]$.

נסיק כי לכל קשת $(u \to v) \in E$ מתקיים כי $f[u] > f[v]$, כלומר u יופיע ב-L לפני v ולכן סידור הצמתים המתקבל הוא אכן מיון טופולוגי. □

תרגיל

שורש בגרף מכוון הוא צומת ממנו יש מסלול (מכוון) לכל צומת אחר בגרף.

א. הציעו אלגוריתם בסיבוכיות $O(V+E)$ שבהינתן גרף מכוון $G=(V,E)$ מוצא שורש ב-G או מודיע שאין כזה. הוכיחו נכונות ונתחו סיבוכיות.

ב. הציעו אלגוריתם בסיבוכיות $O(V+E)$ שבהינתן גרף מכוון $G=(V,E)$ מוצא את כל השורשים ב-G. הוכיחו נכונות ונתחו סיבוכיות.

---

[3] הסימון $O(V+E)$ למעשה אינו מדוייק מאחר ו-V ו-E הינן קבוצות. אם רוצים לדייק יש לרשום $O(|V|+|E|)$, אך אנו נוותר לעיתים על $|\bullet|$ לצורך פשטות ומכיוון שהכוונה ברורה. סימון זה מקובל גם ב-CLRS.

<u>פתרון</u>

א. נריץ DFS על G החל מצומת כלשהו. DFS מחזיר יער מכוון. נסמן ב- r את השורש של העץ האחרון ביער, כלומר הצומת עם זמן הסיום הגדול ביותר. נריץ DFS שוב על G, הפעם החל מ- r. אם בהרצה זו התקבל עץ יחיד (ולא יער) אז נכריז כי r הינו שורש של G. אחרת, נכריז כי אין שורש.

סיבוכיות האלגוריתם היא כמובן $O(V+E)$ כנדרש, מאחר וביצענו שתי הרצות DFS.

<u>**נכונות**</u>: קל לראות כי צומת בגרף הוא שורש אמ"ם הרצת DFS החל ממנו מחזירה עץ ולא יער: אם v הוא שורש אז קיים מסלול מכוון ממנו לכל צומת אחר u. לכן, בתחילת ריצת ה-DFS קיים מסלול לבן מ- v ל- u, לכן u יהיה צאצא של v, כלומר בעץ ה-DFS של v. ולהיפך, אם הרצת DFS מצומת v מחזירה עץ יחיד, אז v שורש העץ ולכן הוא גם שורש בגרף. על כן, האלגוריתם מכריז על r כשורש אמ"ם r אכן שורש.

נניח כעת בשלילה שהאלגוריתם הכריז כי לא קיים שורש, אך קיים בגרף שורש $v \neq r$. אם v שייך לעץ ש- r שורשו (ביחס לריצת ה-DFS הראשונה), אזי קיים מסלול מ- r ל- v ולכן קיים מסלול מ- r (דרך v) לכל צומת אחר, כלומר r שורש. אחרת, v שייך לעץ אחר. נסתכל על המסלול מ- v ל- r ונסמן ב- u את הצומת הראשון על המסלול ששייך לעץ ה-DFS ש- r שורשו (שוב, ביחס לריצת ה-DFS הראשונה). נסמן ב- w את הצומת הקודם ל- u על המסלול. כאשר הקשת $(w \to u)$ נבחנת ע"י ה-DFS (הראשון) u עדיין לא התגלה (כאמור הוא שייך לעץ ה-DFS האחרון שנוצר), לכן הקשת תהיה קשת עץ ו- u יהיה באותו עץ DFS כמו w – סתירה.

ב. על מנת למצוא את כל השורשים מספיק למצוא שורש אחד ולאחר מכן למצוא את הצמתים שמהם ניתן להגיע לאותו שורש. נוכיח זאת באופן פורמלי.

טענה: אם צומת v הינו שורש, אזי צומת u הינו שורש אמ"ם קיים מסלול מ- u ל- v.

הוכחה: נניח שקיים מסלול מ- u ל- v ויהא w צומת בגרף. אזי קיים מסלול מ- v ל- w (כי v שורש), ולכן קיים מסלול מ- u ל- w (דרך v). מצד שני, אם לא קיים מסלול מ- u ל- v אזי ברור כי u אינו שורש (לא ניתן להגיע ממנו ל- v).

נסמן ב- G^t את הגרף $G^t=(V,E^t)$, כך ש- $E^t=\{(u,v)\,|\,(v,u)\in E\}$. כלומר, זה הגרף המתקבל מהפיכת כיווני הקשתות ב- G. ברור כי לכל זוג צמתים $u,v\in V$ קיים מסלול מ- u ל- v ב- G אמ"ם קיים מסלול מ- v ל- u ב- G^t.

מסקנה: על מנת לחשב את כל השורשים ב- G, נחפש שורש אחד בעזרת האלגוריתם מסעיף א'. אם לא מצאנו כזה, אז אין שורשים בגרף. אחרת, יהא r השורש שנמצא. ניצור את הגרף G^t ונריץ DFS עליו החל מ- r. כל צאצאיו של r הם בדיוק שורשי G.

סיבוכיות האלגוריתם היא $O(V+E)$, כנדרש, מאחר והאלגוריתם מבצע 3 ריצות DFS לכל היותר.