הרצאה 1

הקדמה

חיפוש בגרפים, BFS ו-DFS

הקדמה

> אלגוריתם הוא דרך שיטתית (כלומר כזו שצעדיה מוגדרים היטב) לביצוע של משימה מסוימת, במספר סופי של צעדים.

— ויקיפדיה

המושג אלגוריתם אינו חדש עבורנו, ראינו ומימשנו כבר אלגוריתמים בקורס מבוא למדעי המחשב, מבוא לתכנות מערכות ומבני נתונים.

חשיבות הקורס

בתעשייה - בעיות שמצריכות פתרון אלגוריתמי צצות במגוון תחומים. על פי glassdoor, מפתח אלגוריתמים מרוויח 20 אחוז יותר מאשר מהנדס תוכנה.
במחקר האקדמי - חלק עיקרי של המחקר האקדמי הוא בפיתוח וניתוח אלגוריתמים.

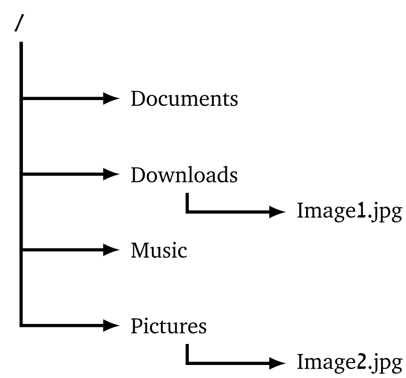
חומר הקורס

בקורס נלמד מגוון אלגוריתמים שעל פי רוב נחשבים לבסיס בתחום האלגוריתמים. הרוב המוחלט של האלגוריתמים שנלמד בקורס הם אלגוריתמים על גרפים וזאת מכיוון שגרפים הוכיחו את עצמם ככלי מאוד חזק בייצוג מגוון רחב של בעיות. לחלק מהאלגוריתמים שימושים ברורים (מסלול קצר ביותר, עצי הופמן), חלק מהאלגוריתמים מהווים פתרון למגוון גדול של בעיות שניתנות לייצוג בצורה מסוימת (זרימה), וחלק מהאלגוריתמים מהווים בסיס לפיתוח אלגוריתמים מורכבים יותר (BFS, DFS, עץ פורש). מעבר לזה נלמד טכניקות (פשטות יחסית) כלליות לפיתוח אלגוריתמים.

אלגוריתמי חיפוש בגרפים

דוגמה 1 (קבצים)**.** רוצים למצוא (ולהדפיס) את כל קבצי התמונות ששמורות על הכונן הקשיח.

למשל עבור:



דוגמה 2 (מבוך)**.** נתון מבוך, נקודת התחלה ונקודת סוף ורוצים למצוא מסלול מנקודת ההתחלה לנקודת הסיום.

דוגמה 3 (פאזל הזזה)**.** נתון לוח משחק בגודל $n \times m$ על הלוח $nm - 1$ חלקים ממוספרים מ-1 עד $nm - 1$ ומשבצת ריקה. נתון סידור ראשוני של החלקים ואנו רוצים לסדר את החלקים לפי הסדר כך שבכל שלב מותר לנו להזיז את אחד החלקים ששכנים למשבצת הריקה אל המשבצת הריקה.

למשל עבור $n = m = 3$:

1	3	6
8	4	
5	2	7

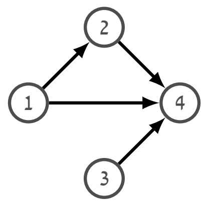
נראה בהמשך שאפשר לייצג כל אחת מהבעיות הנ"ל באמצעות גרף, ומציאת הפתרון לכל אחת מהבעיות מבוסס על סריקה של הגרף המתאים. כיצד עלינו לסרוק כל אחד מהגרפים המתאימים?

ייצוג גרפים

קיימים שני ייצוגים סטנדרטים של גרפים (מכוונים או לא):

1. על ידי מטריצת שכנויות
2. על ידי רשימת שכנויות

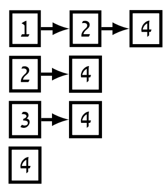
אם לא מצוין אחרת, נניח שהגרף מיוצג על ידי רשימת שכנויות.
למשל את הגרף:



ניתן לייצג על ידי המטריצה

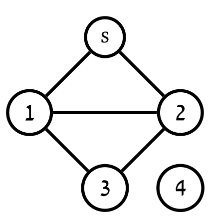
$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

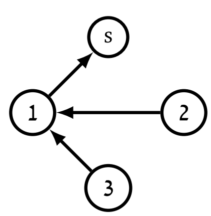
וגם על ידי רשימת שכנויות:



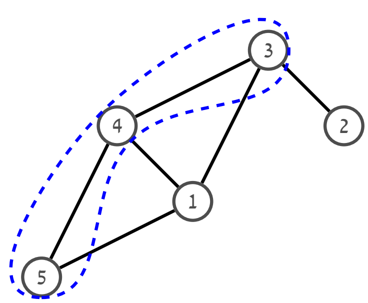
אלגוריתם כללי

קלט: גרף $G=(V,E)$ (מכוון או לא) וצומת מקור s.
פלט: למצוא עץ עם שורש s, $T=(U,F)$, $U\subseteq V$, $F\subseteq E$ כך ש-U היא קבוצת הצמתים שישיגים מ-s. בנוסף, לכל צומת $u\in U$ נרצה ש-$p(u)$ יצביע לאבא של u.
למשל:





הגדרה 1 (חתך). חתך בגרף, $G=(V,E)$, הוא תת קבוצה של צמתים. $S\subseteq V$ נאמר שקשת $uv\in E$ חוצה את החתך S אם $|\{u,v\}\cap S|=1$.



1. אתחול: $T \leftarrow \emptyset$, $U \leftarrow \{s\}$ ולכל $v \in V$ מציבים $p(v) \leftarrow \text{nil}$
2. כל עוד יש קשת uv שחוצה את U ($u \in U$)
   (א) $U \leftarrow U \cup \{v\}$, $T \leftarrow T \cup \{uv\}$, $p(v) \leftarrow u$

טענה 1. בסיום ריצת האלגוריתם U מכילה את כל הצמתים הישיגים מ-s

הוכחה. בשלילה, בוחרים מסלול מ-s לצומת v שלא נכנס ל-U ומסתכלים על הצומת הראשון במסלול שלא נכנס. □

טענה 2. בכל שלב בריצת האלגוריתם T עץ קשיר. בנוסף המסלול מצומת u ל-s הוא שרשור של הקשת $(u, p(u))$ והמסלול מ-$p(u)$ ל-s

הוכחה. באינדוקציה על צעד האלגוריתם. □

חיפוש לרוחב - Breadth First Search (BFS)

הגדרה 2 (מרחק)**.** בהינתן גרף $G = (V, E)$, נגדיר את המרחק בין שני צמתים $u, v \in V$, ונסמנו $dist_G(u, v)$, כמספר הקשתות המינימלי במסלול מ-u ל-v.

הערה: כאשר ברור על איזה גרף מדובר נסתפק בסימון $dist(u, v)$.

קלט: גרף G (מכוון או לא) וצומת מקור s.
פלט: למצוא עץ עם שורש s, $T = (U, F)$, $U \subseteq V$, $F \subseteq E$ כך ש-U היא קבוצת הצמתים שישיגים מ-s. בנוסף, לכל צומת $u \in U$ מתקיים $dist_T(s, u) = dist_G(s, u)$, $d(u) = dist(s, u)$ ו-$p(u)$ יצביע לאבא של u.
למשל:

1. אתחול: $F \leftarrow \emptyset$, $U \leftarrow \{s\}$, לכל $v \in V$ מציבים $p(v) \leftarrow nil, d(v) \leftarrow \infty$, $d(s) \leftarrow 0$
2. כל עוד ישנה קשת uv שחוצה את U ($u \in U$) בחר קשת עם $d(u)$ מינימלי
   (א) $U \leftarrow U \cup \{v\}, F \leftarrow F \cup \{uv\}$
   (ב) $p(v) = u$
   (ג) $d(v) = d(u) + 1$

BFS הוא מקרה פרטי של האלגוריתם הכללי.

טענה 3. לכל $v \in V$ מתקיים $d(v) \geq dist_G(s, v)$

הוכחה. באינדוקציה על צעד האלגוריתם □

טענה 4. לכל $v \in V$ מתקיים $d(v) \leq dist_G(s, v)$

הוכחה. באינדוקציה על צעד האלגוריתם □

טענה 5. לכל $v \in V$ מתקיים $d(v) = dist_T(s, v)$

הוכחה. באינדוקציה על צעד האלגוריתם □

משפט 1. לכל $v \in V$ מתקיים $d(v) = dist_T(s, v) = dist_G(s, v)$

מהו זמן הריצה של BFS ? קשה להגיד כי לא הגדרנו כיצד מתבצעת הבדיקה בשלב 2 של האלגוריתם.

חיפוש לרוחב - מימוש באמצעות תור

ניתן לממש BFS על ידי תור באופן הבא:

1. אתחול: $U \leftarrow \{s\}, F \leftarrow \emptyset$, לכל $v \in V$ מציבים $p(v) \leftarrow nil, d(v) \leftarrow \infty$, $d(s) \leftarrow 0, Q \leftarrow (s)$
2. כל עוד התור לא ריק
   (א) $u \leftarrow Q.pop()$
   (ב) כל עוד ישנה קשת uv שחוצה את U ($u \in U$)
      i. $U \leftarrow U \cup \{v\}, F \leftarrow F \cup \{uv\}$
      ii. $p(v) = u$
      iii. $d(v) \leftarrow d(u) + 1$
      iv. $Q.push(v)$

נשים לב שבמימוש הנ"ל כל צומת נכנסת ויוצאת מהתור לכל היותר פעם אחת ובנוסף כל קשת נבדקת לכל היותר פעם אחת ולכן זמן הריצה הוא $O(|V| + |E|)$ נשאר להוכיח שזהו אכן מימוש של BFS.

הגדרה 3 (צומת גבולי)**.** בהינתן גרף $G = (V, E)$ וחתך $U \subseteq V$ צומת $u \in U$ יקרא גבולי אם קיימת קשת uv שחוצה את U.

טענה 6. בכל שלב בריצת האלגוריתם התור מכיל את כל הצמתים הגבוליים

הוכחה. באינדוקציה על צעד האלגוריתם □

טענה 7. התור מונוטוני לא יורד בהתייחס לערכי d

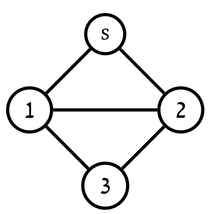
הוכחה. נוכיח טענה חזקה יותר באינדוקציה על צעד האלגוריתם: התור מונוטוני לא יורד וגם $|d(u) - d(v)| \leq 1$ לכל שני צמתים שבתור □

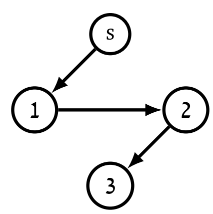
מסקנה 1. זהו אכן מימוש של BFS

חיפוש לעומק - Depth First Search (DFS)

נגדיר $\mathrm{dfs}(U, T, u)$

1. אם קיימת קשת uv שחוצה את U אז
   (א) $U \leftarrow U \cup \{v\}, T \leftarrow T \cup \{uv\}, p(v) \leftarrow u$
   (ב) $\mathrm{dfs}(U, T, v)$





סיכום

דוגמה 4 (פאזל הזזה)**.** נתון לוח משחק בגודל $n \times m$ על הלוח $nm - 1$ חלקים ממוספרים מ-1 עד $nm - 1$ ומשבצת ריקה. נתון סידור ראשוני של החלקים ואנו רוצים לסדר את החלקים לפי הסדר כך שבכל שלב מותר לנו להזיז את אחד החלקים ששכנים למשבצת הריקה אל המשבצת הריקה.

למשל עבור $n = m = 3$:

1	3	6
8	4	
5	2	7

הציעו אלגוריתם לפתרון הבעיה.