

מבני נתונים - תרגיל 8

תאריך הגשה: 01.06.2023

בתרגיל זה נעסוק באלגוריתמים BFS ו-DFS אותם ראינו בהרצאה ובתרגול. אלא אם נאמר אחרת, התייחסות לאלגוריתמים שראינו בכיתה היא עבור הגרסה הספציפית אותה ראינו. כל הגרפים מאוחסנים כרשימת שכנויות (adjacency-list) אלא אם יוגדר אחרת. כשאתם מתבקשים להציג אלגוריתם יש לנתח בקווים כלליים זמן ריצה ולהסביר באופן כללי למה האלגוריתם עובד. אם תדרש פורמאליות - יהיה כתוב מפורשות "הוכיחו". ניתן לתאר אלגוריתמים באיזו דרך שתמצאו ובתנאי שהיא ברורה ומובנת. זכרו - לפעמים (אך לא תמיד) תיאור במילים שווה לאלף פסודוקודאים.

חלק א (26 נקודות)

שאלות קצרות על גרפים, BFS ו-DFS

א. [6 נק'] מה זמן הריצה של אלגוריתם BFS אם הקלט שלו הוא גרף המיוצג ע"י מטריצת שכנויות? הסתכלו על הפסודו קוד של האלגוריתם הנתון בתרגול והסבירו מהי הנקודה הספציפית בפסודו קוד בו יידרש שינוי להתמודדות עם ייצוג זה של הגרף.

ב. [6 נק'] נתון גרף לא מכוון: $G = (V, E)$. תזכורת מדיסקרטית:

$$|E| \begin{cases} \geq |V| - 1 & G \text{ קשיר} \\ = |V| - 1 & G \text{ עץ} \\ \leq |V| - 1 & G \text{ חסר מעגלים} \end{cases}$$

נריץ את הגרסה של אלגוריתם DFS שעוברת על כל רכיבי הקשירות בגרף הנתון G . אנו יודעים כי זמן ריצה של אלגוריתם זה הוא $O(|E| + |V|)$. יחד עם זאת, זמן ריצה זה איננו (בהכרח) שקול ל- $O(|E|)$ או ל- $O(|V|)$. לכל אחד מסוגי הגרפים הבאים ציינו האם זמן הריצה של אלגוריתם DFS שקול ל- $O(|E|)$ או ל- $O(|V|)$ (או שניהם אם הם אותו סדר גודל):

1. G קשיר

2. G עץ (=קשיר וחסר מעגלים)

3. G חסר מעגלים (= G יער)

ג. [8 נק'] נתון גרף G כלשהו בעל 5 קודקודים. סטודנטית בקורס מבני נתונים הריצה את אלגוריתם DFS על הגרף וקיבלה את זמני הכניסה (pre) הבאים אותם היא מיינה לפי סדר עולה: 1, 2, 4, 5, 6. רשמו מהם זמני היציאה (post) של כל אחד מהקודקודים (כלומר לכל זמן כניסה של כל אחד מהקודקודים התאימו את זמן היציאה שמתאם לקודקוד הזה) ובנוסף ציירו את העץ הנוצר מריצת DFS על גרף G .

ד. [6 נק'] ציינו ל-3 הטענות הבאות אם הן נכונות או לא נכונות. אם התשובה אינה נכונה, ספקו דוגמה נגדית. אם היא נכונה, הסבירו בקצרה ובמילים את תשובתכם:

1. יהי T עץ הרוחב המתקבל מהרצת BFS על קודקוד מקור s בגרף קשיר ולא מכוון G . אזי, מתקיים ש- T מכיל בדיוק את כל הקודקודים בגרף.

2. יהי T עץ הרוחב המתקבל מהרצת BFS על קודקוד מקור s בגרף מכוון G ובו ישנה קשת בכיוון אחד בלבד בין כל שני קודקודים. אזי, מתקיים ש- T מכיל בדיוק את כל הקודקודים בגרף.

3. יהי T עץ העומק המתקבל מהרצת DFS על קודקוד מקור s בגרף G שהוא גרף של מעגל לא מכוון עם $|V| \geq 3$ קודקודים. אזי, T בהכרח אינו מכיל את כל המסלולים הקצרים ביותר מקודקוד המקור s לכל הקודקודים האחרים בגרף.

חלק ב (42 נקודות)

ויקיפדיה

[הערך הראשון](#) בויקיפדיה העברית נכתב ב-2003. מאז נכתבו עוד 334,287 ערכים (!) בעברית (נכון לכתיבת שורות אלו). ניתן לחשוב על ויקיפדיה כעל גרף ענק כך שכל ערך בו הוא קודקוד והצלעות בין הקודקודים אלו הערכים שמוזכרים בערך אחר ([המחשה](#)). שימו לב שויקיפדיה הוא גרף מכוון, לדוגמה "כלב" לא יופיע ב"טורפים" אך ההפך כן.

א. [10 נק'] בחרו ערך מויקיפדיה ובצעו עליו BFS ו-DFS כפי שראינו בכיתה. כל ערך הוא קודקוד ונניח שהשכנים היחידים שלו הם שני הערכים הראשונים שמוזכרים בו. בנוסף, בשביל שנסיים את התואר בזמן, נניח שלכל קודקוד שמרחקו 2 מהשורש אין צלעות היוצאות ממנו כלל. ציירו את העצים שנוצרו. הוסיפו לגרף של DFS זמני כניסה ויציאה בכל קודקוד.

ב. [10 נק'] לצערנו לא מצאנו נתונים על מספר הצלעות שיש לויקיפדיה. בהנחה שהגרף מאוחסן בויקיפדיה כרשימת שכנויות (adjacency-list), הציעו אלגוריתם יעיל כל שתוכלו המוצא את מספר הצלעות בגרף.

ג. [6 נק'] ויקיפדיה העברית שוקלת בסה"כ [5.7Gb](#). במטרה לחסוך עוד יותר מקום הוטל עליכם להחליט איזה מבנה נתונים הוא המתאים ביותר לאחסון הגרף של ויקיפדיה. מבנה זה ישמש את ויקיפדיה לשנים ארוכות. באיזה מבנה מאלו שראינו בכיתה תבחרו? נמקו.

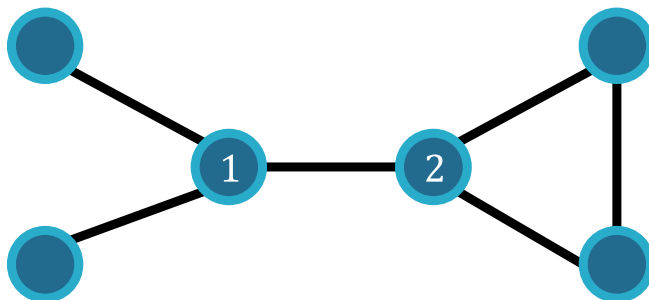
ד. [16 נק'] האתר [naraview](#) הוא משחק ישראלי שמאתגר משתתפים שונים לנווט בין שני ערכים שונים בויקיפדיה. כל המשתתפים מתחילים מאותו הערך והמשתתף שמספר הלחיצות שלו (מספר הקודקודים אותם עבר) הוא הקטן ביותר עד להגעה לערך השני - מנצח. יכול להיות שלא קיים מעבר מערך אחד לערך אחר גם לאחר אינסוף לחיצות. נמקו מדוע והציעו אלגוריתם יעיל ככל שתוכלו שעבור מבנה הנתונים שבניתם בסעיף קודם מוצא את מספר הזוגות שלא ישיגים אחד עבור השני. נניח את ההנחה הסבירה שתמיד נוכל לחזור חזרה לערך בו ביקרנו באמצעות מספר לחיצות סופי.

הערה: הקישור של naraview מוביל אל משחק בין הערכים BFS ל-**כלב ים נזירי מצוי**. בפתרון התרגיל נפרסם את הגרף הגדול שיצרו כל הסטודנטים בקורס (כל המסלולים שעברו בהם). הזוכה במסלול הקצר ביותר יזכה בתואר היוקרתי "מלכת ה-BFS" ומקום של כבוד בפתרון התרגיל. הקישור תקף ליומיים לכן יתבצעו שלושה סבבים ונפרסם את הטוב מביניהם (בכל פעם נעדכן את הקישור). כמו כן, פרס הנחישות או בשמו השני "פרס ה- $\Theta(|V|^2)$ " יוענק למי שיצרה את המסלול הארוך ביותר.

חלק ג (32 נקודות)

שאלות אלגוריתמיות

א. [16 נק'] לעיתים בגרפים יכול להיווצר "צוואר בקבוק" בו מעבר בין שתי קבוצות קודקודים מתרכז בצלע אחת. נרצה למצוא קודקודים שמהווים כניסה לצוואר בקבוק. ובאופן יותר פורמאלי - יהי G גרף לא מכוון. נגדיר "צלע בקבוק" ככזו שהסרה שלה תגרום לרכיב הקשירות בו היא נמצאת להתפצל ל-2 רכיבי קשירות שונים. מצאו אלגוריתם שעבור קודקוד s מסוים מוצא את מספר צלעות הבקבוק של אותו קודקוד.



איור 0.1: בדוגמה זו לקודקוד 1 יש 3 "צלעות בקבוק" ול-2 יש 1.

ב. [16 נק'] הגדרה: גרף לא מכוון G נקרא דו-צדדי אם ניתן לחלק את הקודקודים בו ל-2 קבוצות זרות כך שלא קיימת קשת בין שני קודקודים השייכים לאותה קבוצה. תזכורת מדיסקרטית (ניתן להשתמש בטענה זו לפתרון השאלה אך קיימים עוד הרבה פתרונות בלעדית): גרף הוא דו-צדדי אם אין בו מעגל באורך אי-זוגי. הציעו אלגוריתם יעיל ככל שתוכלו, המקבל גרף לא מכוון G (לאו דווקא קשיר) ומחזיר True במידה והגרף הוא דו צדדי ו-False אחרת. הסבירו את נכונות האלגוריתם (אין צורך בהוכחת נכונות פורמאלית).