אלגוריתמים קומבינטוריים
סיכומים של תרגילי כיתה מסמסטרים קודמים בנושא
חיפוש לרוחב BFS
(Breadth First Search)

1. נתון גרף (לא מכוון או מכוון) $G(V,E)$ וצומת מקור s.

2. הרעיון של חיפוש לרוחב הוא שנבקר בצמתים שניתן להגיע אליהן מ- s על ידי מסלול (מכוון) ב- G, ושהביקורים יהיו מסודרים לפי המרחק (לפי אורך המסלול שהוא מספר הקשתות במסלול) מ- s. במשך האלגוריתם נגדיר את המרחק של כל צומת מ- s ואת אביו בעץ ה-BFS, כלומר הקודם לו במסלול קצר ביותר מהשורש s.

3. זמן הריצה של BFS הוא $O(|V|+|E|)$. כאשר $|E|=\Omega(|V|)$ (ובפרט כאשר G קשיר) הסיבוכיות הוא $O(|E|)$. (הוכחה בהמשך.)

4. הגדרה: יהי $G(V,E)$ גרף. עבור $u,v\in V$ נסמן ב- $d(u,v)$ את המרחק מ- u ל- v ב- G (אורך המסלול הקצר ביותר מ- u ל- v ב- G).

5. הגדרה: יהי $G(V,E)$ גרף קשיר. הקוטר (diameter) של G מוגדרת להיות

$$\operatorname{diam}(G)=\max_{u,v\in V}\{d(u,v)\}$$

(האורך המקסימלי של מסלול קצר ביותר בין זוג צמתים).

6. הערה: על ידי הרצת BFS מכל צומת בגרף, אפשר למצוא לכל צומת u את הצומת הכי רחוק ממנו v_u ואת מרחקם $d(u,v_u)$. על ידי מציאת המקסימום $\max_{u\in V}\{d(u,v_u)\}$ נקבל את $\operatorname{diam}(G)$.
מסקנה: בגרף כללי ניתן לחשב את הקוטר בזמן $O(|V||E|)$.

7. תרגיל: תן אלגוריתם לחשב את הקוטר של עץ, T, בסיבוכיות $O(|E|)$.

8. למה 1: בעץ יש מסלול יחיד בין כל זוג של צמתים.

9. הוכחת למה 1: נניח בשלילה שיש שני מסלולים שונים P_1 ו- P_2 מ- u ל- v ב- T. יהי u' הצומת האחרון המשותף לתחילת שני המסלולים (אולי $u'=u$). מכיון שכל אחד מבין המסלולים המקוריים מגיע ל- v יש גם צומת v' משותף ראשון בהמשך המסלולים. בכך מצאנו מעגל ב-T בסתירה להיותו עץ.

10. למה 2: יהי P מסלול מ- a ל- b ($a\neq b$) בעץ T שאינו ניתן להארכה בצד של b. אזי b עלה.

11. הוכחת למה 2: ל- b מגיע הקשת האחרון במסלול. נניח בשלילה שיש עוד קשת e הנוגע ב- b. אם הקצה השני של e הוא במסלול P אז נסגר מעגל - בסתירה לזאת ש- T עץ. אחרת, היה ניתן להמשיך את המסלול על ידי שימוש בקשת e בסתירה להנחה ש- P אינו ניתן להארכה בצד של b. לכן אין קשת נוספת על b, כלומר ש- b עלה.

12. למה 3: יהי P מסלול מ- a ל- b בעץ T. אזי לכל $v\in V\setminus V(P)$ קיימת צומת $x_v\in V(P)$ שדרכו עובר כל מסלול בין צומת ב-P ו- v.

13. הוכחת למה 3: אחרת יהיה מעגל בגרף - בסתירה להיותו עץ. (השלם את הפרטים!)

14. טענה: יהי T עץ ו- s צומת כלשהי ב- T. יהי t צומת במרחק המקסימלי מ- s. אזי t הוא הקצה של מסלול באורך $\operatorname{diam}(T)$.

15. נדחה בינתיים את הוכחת הטענה. קודם נפתח אלגוריתם לפתור את התרגיל:

```
TREE-DIAMETER(T)
1  s ← some vertex in T   ▷ s arbitrary.
2  t ← BFS-LAST(G, s)   ▷ Find t.
3  D ← BFS-MAX(G, t)   ▷ Find max distance from t.
4  return (D)
```

```
BFS(G, s)   ▷ No return at end.
BFS-MAX(G, s)   ▷ Last row returns max distance.
BFS-LAST(G, s)   ▷ Last row returns a vertex.
1     for each vertex u ∈ V \ {s}
2         do color[u] ← WHITE
3            d[u] ← ∞
4            π[u] ← NIL
5     color[s] ← GREY
6     d[s] ← 0
7     π[s] ← NIL
7a    D ← 0   ▷ For DFS-MAX.
7b    last ← s   ▷ FOR DFS-LAST.
8     Q ← {s}   ▷ Q is a queue (FIFO).
9     while Q ≠ ∅
10        do u ← head[Q]
11           for each v ∈ Adj[u]
12               do if color[v] = WHITE   ▷ Newly discovered.
13                   then color[v] ← GREY
14                        d[v] ← d[u] + 1
14a                       D ← d[v]   ▷ For BFS-MAX.
14b                       last ← v   ▷ For BFS-LAST.
15                        π[v] ← u
16                        ENQUEUE(Q, v)
17           DEQUEUE(Q)
18           color[u] ← BLACK
19a   return (D)   ▷ For BFS-MAX.
19b   return (last)   ▷ For BFS-LAST.
```

נשים לב שהצומת אחרון שיתגלה על ידי BFS הוא במרחק מקסימלי מ- s, והוא הערך האחרון ש- $last$ מקבל. המרחק מ- s ל- $last$ הוא הערך האחרון ש- D מקבל. לכן אם הטענה נכונה אז האלגוריתם יחזיר את קוטר העץ T.

16. ניתוח סיבוכיות לכל שלושת הגירסאות של BFS:

- האיתחול בשורות 1 עד 8 דורש זמן $O(|V|)$.
- לצומת מסיום u מהתור הלולאה משורות 11 ל- 16 עובר על כל $\text{outdeg}(u)$ השכנים של u ולכל אחד דרוש $O(1)$ זמן טיפול. סה"כ $O(\text{outdeg}(u))$.
- כל צומת נכנס לתור לכל היותר פעם אחד. (רק כאשר נתגלה בפעם הראשון, אך ייתכן שלא יתגלה - אם אין מסלול מ- s אליו.) כל צומת שנכנס לתור יוצא מהתור בסוף טיפולו - ולכן מופיע כ- u בשורה 10 בדיוק פעם אחד. הטיפול ב- u בשורות 10, 17, ו- 18 דורש זמן $O(1)$.

- לכן זמן הריצה של שורות 9 עד 18 הוא

$$\sum_{u \text{ goes through } Q} O(\text{outdeg}) \leq O(\sum_{u \in V} \text{outdeg}) = O(|E|).$$

- החזרת תשובה על ידי שורה 19 דורשת $O(1)$ זמן.
- סיבוכיות כוללת: $O(|V| + |E|)$.

17. מסקנה: האלגוריתם TREE-DIAM רץ בזמן $O(2(|V| + |E|)) = O(|E|) = O(|V|)$ כי T עץ.

18. הוכחת הטענה: (מומלץ לצייר את הציורים המתאימים במשך ההוכחה.)

יהי P מסלול מ- a ל- b באורך $\text{diam}(T)$ ב- T. אם $a = b$ אז $|V| = 1$ והטענה טרוויאלית. לכן נניח ש- $a \neq b$. לכן, לפי למה 2, a ו- b עלים.

אם $t \in \{a, b\}$ אז P הוא מסלול כנדרש. אחרת, מכיון ש- t הוא צומת במרחק מקסימלי מ- s, t הוא עלה לפי למה 2; ואז t לא מופיע כצומת פנימי במסלול P. לכן קיים צומת $x_t \in V(P)$ כמתואר בלמה 3. ($x_t \notin \{a, b\}$ כי a ו- b עלים.)

נסמן $\alpha = d(a, x_t), \beta = d(x_t, b)$. אזי $\alpha + \beta = \text{diam}(T)$.

נניח ש- s מופיע כצומת פנימית במסלול מ- t ל- x_t. (במקרה הזאת קל לראות כי $x_s = x_t$ וכן $d(s, x_s) > 0$.) נשתמש בעובדה ש- $d(s, a) \leq d(s, t)$ (t במרחק מקסימלי מ- s) לקבל:

$$\begin{aligned} d(t, b) &= d(t, s) + d(s, x_t) + d(x_t, b) \\ &\geq d(s, a) + d(s, x_s) + \beta \\ &= [d(s, x_s) + d(x_s, a)] + d(s, x_s) + \beta \\ &= \alpha + \beta + 2d(s, x_s) \\ &> \alpha + \beta \\ &= \text{diam}(T) \end{aligned}$$

וזו סתירה. לכן s אינו מופיע כצומת פנימית במסלול מ- t ל- x_t.

במקרה ש- s מופיע כצומת במסלול P נגדיר את $x_s = s$ אחרת נגדיר את x_s לפי למה 3. נניח בלי הגבלת הכלליות ש- x_s מופיע בין x_t ו- b (על ידי החלפה בין a ו- b לפי הצורך. ייתכן גם ש- $x_s = x_t$.) [צייר!!!] שוב נשתמש בעובדה ש- $d(s, a) \leq d(s, t)$. אזי:

$$d(s, x_s) + d(x_s, a) = d(s, a) \leq d(s, t) = d(s, x_s) + d(x_s, t)$$

ולכן $d(x_s, a) \leq d(x_s, t)$. אזי:

$$d(b, t) = d(b, x_s) + d(x_s, t) \geq d(b, x_s) + d(x_s, a) = d(b, a) = \text{diam}(T) \geq d(b, t)$$

אך ברור ש- $d(b, t) \leq \text{diam}(T)$ לפי הגדרת הקוטר. לכן, $d(b, t) = \text{diam}(T)$, ו- t הוא הקצה של מסלול באורך $\text{diam}(T)$ כנדרש. [סוף הוכחת הטענה.]

19. הגדרנו מהו עץ ה-BFS, העץ המכוון עם שורש s של הקשתות מהסוג $(\pi(v), v) : v \in V \setminus \{s\}$. שמנו לב שזה עץ פורש של רכיב הקשירות המכיל את s (בגרף לא-מכוון בגרף מכוון זה עץ פורש של אוסף הקודקודים שניתן להגיע אליהם מ-s).

20. הערנו על ההבדל בגישת החיפוש לרוחב לעומת החיפוש לעומק.