תרגול 8 – Huffman ומסלולים קלים ביותר

תרגיל 1 – Huffman

הוכח / הפרך כל אחת מהטענות הבאות

1. אם כל תו מופיע בקובץ בתדירות קטנה ממש מ-$1/3$ אזי בהכרח בעץ Huffman לא תהיה מילת קוד מאורך 1.
2. קיים אלגוריתם כיווץ אשר מקטין ממש את גודלו של <u>כל</u> קובץ.

פתרון

סעיף 1

<u>הטענה נכונה</u>. נניח בשלילה כי קיים תו עם מילת קוד מאורך 1. נסמן ב-u את הצומת המתאים למילה זו בעץ Huffman, ונבחין כי משקל u קטן ממש מ-$1/3$. נסמן ב-r את שורש עץ Huffman. מהנחת השלילה ישנה הקשת ru. נביט בבן השני של r בעץ, שנסמנו ב-v. מהנחת התרגיל נובע שהמשקל המתאים ל-v גדול ממש מ-$2/3$. אם כך v הוא צומת פנימי בעץ (שכן לכל עלה יש משקל קטן ממש מ-$1/3$). כצומת פנימי, ל-v יש בדיוק שני ילדים. נבחין כי אחד מהם הוא במשקל גדול ממש מ-$1/3$ (שכן אם משקל שני ילדיו אינו עולה על $1/3$, אז משקלו הוא אמור להיות לכל היותר $2/3$). נסמן צומת זה ב-w, ואת הצומת השני ב-y.

בעת יצירת הצומת v לרשות Huffman היו שלושה צמתים - u,w ו-y. האלגוריתם חיבר את y,w ולכן, מהגדרת האלגוריתם, משקל כל אחד מהם אינו עולה על משקלו של u. אבל הראנו כי משקלו של u קטן ממש מ-$1/3$ ואילו משקלו של w גדול ממש מ-$1/3$. סתירה.

סעיף 2

הטענה אינה נכונה. אנו נוכיח טענה חזקה יותר. נוכיח שלכל אלגוריתם כיווץ ולכל n טבעי, קיימת מילה $x \in \{0,1\}^n$ שהאלגוריתם אינו מכווץ (כלומר, לא רק שלכל אלגוריתם יש מילה שהאלגוריתם אינו מצליח לכווץ, אלא שיש מילה כזו מכל אורך). כל אלגוריתם כיווץ משרה פונקציה $f:\{0,1\}^* \to \{0,1\}^*$ כך ש-x מכווץ למחרוזת $f(x)$. אנו לא יודעים דבר על f, שכן היא תלויה באלגוריתם, פרט לכך שהיא חח"ע (אחרת לא ניתן היה לשחזר את הקובץ). יהא n כלשהו. נביט בפונקציה f מצומצמת למחרוזות מאורך n. כלומר נביט ב-$f':\{0,1\}^n \to \{0,1\}^*$ המזדהה עם f על $\{0,1\}^n$. נניח בשלילה ש-f' מכווצת את כל המחרוזות ב-$\{0,1\}^n$. כלומר f' היא פונקציה חח"ע מהצורה

$$f':\{0,1\}^n \to \bigcup_{k=0}^{n-1}\{0,1\}^k$$

לכן, מכיוון ש-f' חח"ע

$$\left|\{0,1\}^n\right| \le \left|\bigcup_{k=0}^{n-1}\{0,1\}^k\right|$$

אבל

$$\left|\{0,1\}^n\right| = 2^n$$

$$\left|\bigcup_{k=0}^{n-1}\{0,1\}^k\right| = \sum_{k=0}^{n-1} 2^k = 2^n - 1$$

סתירה.

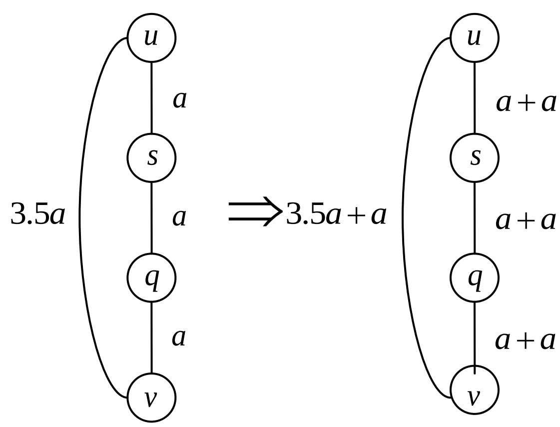
תרגיל 2

נתון גרף קשיר ולא מכוון $G=(V,E)$ וכן פונקציית משקל $w: E \to R$.

1. נגדיר פונקציית משקל חדשה $c: E \to R$ באופן הבא: $c(e) = w(e) + a$ (a מספר אי שלילי כלשהו). הוכיחו/הפריכו: אם P הינו מסלול קל ביותר מ-u ל-v תחת w אזי P מסלול קל ביותר מ-u ל-v תחת c.
2. נגדיר פונקציית משקל חדשה $c: E \to R$ באופן הבא: $c(e) = a \cdot w(e)$ (a מספר חיובי כלשהו). הוכיחו/הפריכו: אם P הינו מסלול קל ביותר מ-u ל-v תחת w אזי P מסלול קל ביותר מ-u ל-v תחת c.

פתרון

1. הטענה אינה נכונה. דוגמה נגדית:



2. הטענה נכונה. נשים לב כי משקל כל מסלול בגרף מוכפל פי a. נניח בשלילה כי קיים מסלול Q מ-u ל-v כך שמתקיים: $c(Q) < c(P)$. נקבל:

$$\sum_{e\in Q} a\cdot w(e) = \sum_{e\in Q} c(e) = c(Q) < c(P) = \sum_{e\in P} c(e) = \sum_{e\in P} a\cdot w(e).$$

מכאן נסיק:

$$a \cdot w(Q) = a\sum_{e\in Q} w(e) = \sum_{e\in Q} a \cdot w(e) < \sum_{e\in P} a \cdot w(e) = a\sum_{e\in P} w(e) = a \cdot w(P)$$

ולכן בהכרח $w(Q) < w(P)$. סתירה.

תרגיל 3

נתון גרף מכוון וקשיר $G=(V,E)$ עם משקלים שלמים חיוביים על הקשתות. הצע אלגוריתם המוצא עבור כל צומת $v \in V$ את משקל המסלול הקל קצר ביותר מ- s ל- v . כלומר, מבין המסלולים הקלים ביותר מ- s ל v , מסלול קל קצר הינו המסלול בעל המספר המינימלי של קשתות.

פתרון

נרצה לתת עדיפות למסלולים קצרים יותר. לשם כך נגדיר פונקצית משקל חדשה

$$w'(u,v) = |V| \cdot w(u,v) + 1$$

טענה 1

P הוא מסלול קל קצר ביותר לפי פונקצית המשקל w אם"ם P הוא מסלול קל ביותר לפי פונקצית המשקל w' .

הוכחה

ראשית ניקח שני מסלולים פשוטים P,Q כך ש- $w(P) < w(Q)$, אז

$$\mathrm{w}'(\mathrm{Q}) - \mathrm{w}'(\mathrm{P}) = \sum_{(\mathrm{u},\mathrm{v})\in \mathrm{Q}} (|\mathrm{V}| \cdot \mathrm{w}(\mathrm{u},\mathrm{v}) + 1) - \sum_{(\mathrm{u},\mathrm{v})\in \mathrm{P}} (|\mathrm{V}| \cdot \mathrm{w}(\mathrm{u},\mathrm{v}) + 1)$$

$$= (|V| \cdot w(Q) + |Q|) - (|V| \cdot w(P) + |P|) = |V|\big(w(Q) - w(P)\big) + |Q| - |P|$$

היות ו-$|P| \le |V| - 1$ מהיותו מסלול פשוט, וגם $w(Q) - w(P) \ge 1$ שכן המשקלים שלמים, אזי: $w'(Q) - w'(P) \ge |V| - |V| + 1 > 0$ ← $w'(Q) > w'(P)$ כנדרש.

שנית, נבחין כי אם $w(P) = w(Q)$ אז $w'(P) < w'(Q)$ אם"ם $|P| < |Q|$.

$$w'(Q) - w'(P) = \sum_{(u,v)\in Q}(|V| \cdot w(u,v) + 1) - \sum_{(u,v)\in P}(|V| \cdot w(u,v) + 1) =$$
$$|V|w(Q) + |Q| - |V|w(P) - |P| > 0$$

■

טענה 1 מאפשרת לנו לעשות רדוקציה מבעיית מסלול קל קצר ביותר לבעיית מסלול קל ביותר "רגילה" אותו אנו יודעים לפתור. אם כך, יש לנו את האלגוריתם הבא לפתרון בעיית מסלול קל קצר ביותר.

אלגוריתם

<u>קלט</u>: גרף מכוון וקשיר $G=(V,E)$, $s \in V$, ופונקצית משקל $w: E \to N$.

<u>פלט</u>: מסלולים קלים קצרים ביותר מ-s לכל $v \in V$.

1. חשב את פונקצית המשקל w'.
2. החזר מסלולים קלים ביותר מ-s לכל v לפי 'w.

נכונות

נכונות האלגוריתם נובעת מטענה 1.

סיבוכיות

מכיוון שחישוב w' ניתן לעשות בזמן לינארי באורך הקלט, סיבוכיות האלגוריתם היא כסיבוכיות האלגוריתם אותו אנו מפעילים בשלב 2 למציאת מסלולים קלים ביותר. האלגוריתם היעיל ביותר שראינו בקורס (המתאים למשקלים אי-שליליים) רץ בזמן $O(E \cdot logV)$ ולכן זוהי גם סיבוכיות זמן הריצה של האלגוריתם אותו אנו מנתחים.