אלגוריתמים הרצאה 10

תכנון דינאמי - המשך

דוגמה 3:

מסלולים קלים ביותר בין כל זוגות הצמתים בגרף
נתון: גרף מכוון $G = (V, E)$
פונקציית משקל $w : E \to \mathbb{R}$ (נניח שאין מעגלים שליליים)
מטרה: לכל שני צמתים u ו-v, נרצה לחשב את אורך המסלול הקל ביותר ב-G מ-u ל-v.
פתרון פשוט: להריץ Bellman Ford מכל צומת n.
סיבוכיות: $O(|V|^2 \cdot |E|)$.

שאלה:

כיצד להגדיר תתי-בעיות?
צמתי הבנים של מסלול p יהיו כל הצמתים ב-p פרט לצומת ההתחלה וצומת הסיום.
נניח מטעני נוחות, שצמתי הגרף הם $\{v_1, \ldots, v_n\}$.
נגדיר תת-בעיה: לכל זוג צמתים u ו-v ולכל $k = 0, 1, \cdots, n$ ונסמן $\delta_k(i, j)$ להיות משקל מסלול
ב-G מ-v_i-ל-v_J מבין כל המסלולים שצמתי הביניים שלהם לקוחים מתוך $\{v_1, \ldots, v_k\}$.

$$D_k(i,j) = \begin{cases} 0 & k = 0, i = j \\ w_{i,j} & k = 0, i \neq j, (v_i, v_j) \in E \\ \infty & k = 0, i \neq j, (v_i, v_j) \notin E \\ min\{D_{k-1}(i,j), D_{k-1}(i,k) + D_{k-1}(k,j) & \text{otherwise} \end{cases}$$

טענה:

$D_k(i, j) = \delta_k(i, j)$ לכל $1 \leq i, j \leq n$ ולכל $k \in \{0, 1, \ldots, n\}$

- **שאלה:** כיצד לחשבת את ה-D-ים ? (באופן מהיר)
  חישוב כל מטריצה לוקח $O(n^2)$ זמן
  $\Leftarrow$ סה"כ $O(n^3)$

$$\underbrace{\boxed{\begin{matrix}D_0(i,j)\\ (nxn)\end{matrix}}}_{\underbrace{k=0}_{\text{(stop cond.)}}} \Rightarrow \underbrace{\boxed{\begin{matrix}D_1(i,j)\\ (nxn)\end{matrix}}}_{k=1} \Rightarrow \cdots \Rightarrow \underbrace{\boxed{\begin{matrix}D_n(i,j)\\ (nxn)\end{matrix}}}_{k=n}$$

⋆ **הערה:** האלגוריתם צורך $O(n^2)$ במקום.

אלגוריתם זה נקרא Floyd Warshall.

הוכחה:

באינדוקצייה על k.
<u>בסיס:</u> $k=0$ נובע מידי מהגדרת $\delta_0(i,j)$ ו־$D_0(i,j)$
<u>צעד:</u> נקבע שני צמתים v_i ו־v_j ונזכר ש־$\delta_k(i,j)$ הוא אורך המסלול הקל ביותר מ־v_i ל־v_j ב־G שצמתי הביניים שלו לקוחים מתוך $\{v_1,\dots,v_k\}$. יהי p מסלול כזה.
נחלק לשני מקרים:

1. נניח כי v_k נמצא ב־p, נשים לב ש־v_K מופיע פעם אחת $v_i \longrightarrow v_k \longrightarrow v_j$.
ב־p כי אין מעגלים שליליים שמסלול קל ביותר הוא פשוט.
$=\delta_k(i,j)$
סך משקלי הקשתות ב־p=
סך משקלי הקשתות בסיפא מ־v_k ל־ v_j+סך משקל הקשתות ברישא מ־v_iל־v_k
$\delta_k(i,j)=\delta_{k-1}(i,k)+\delta_{k-1}(k,j)$.
נשים לב שסך משקלי הקשתות ברישא של p מ־v_i ל־v_k שווה ל־$\delta_{k-1}(i,k)$
שכן אחרת $\delta_{k-1}(i,k)$ קטן יותר וניתן להחליף את הרישא במסלול טוב יותר מ־v_iל־v_k
שצמתי הביניים שלו מתוך $\{v_i,\dots,v_{k-1}\}$ בסתירה לאופטימליות של k).
באופן דומה, סך משקלי הקשתות של הסיפא של p מ־v_kל־v_J שווה ל־ $\delta_{k-1}(k,j)$.

$$\delta_k(i,j)=\delta_{k-1}(i,j)+\delta_{k-1}(i,j)\underbrace{=}_{\text{induction}}D_{k-1}(i,k)+D_{k-1}(k,j)$$

נראה ש־ $D_{k-1}(i,j)\geq D_{k-1}(i,k)+D_{k-1}(k,j)$

$$D_{k-1}(i,j)\underbrace{=}_{\text{induction}}\delta_{k-1}(i,j)\geq\delta_k(i,j)\underbrace{=}_{\text{proved}}D_{k-1}(i,k)+D_{k-1}(k,j)$$

$\Longleftarrow$ $\delta_k(i,j)=D_k(i,j)$.

2. נניח כי v_K לא נמצא ב־p.

$$\Rightarrow \delta_k(i,j)=\Sigma(\text{weights in p})\underbrace{=}_{v_k\text{not in p}}\delta_{k-1}(i,j)\underbrace{=}_{\text{induction}}D_{k-1}(i,j)$$

נראה ש־$D_{k-1}(i,j)\leq D_{k-1}(i,k)+D_{k-1}(k,j)$

$$D_{k-1}(i,j)+D_{k-1}(k,j)\underbrace{=}_{\text{induction}}\delta_{k-1}(i,k)+\delta_{k-1}(k,j)\underbrace{\geq}_{\text{definition}}\delta_k(i,j)=D_{k-1}(i,j)$$

$$\Longleftarrow \delta_k(i,j)=D_k(i,j)$$

המקרה הנותר הוא שאין מסלול מ־v_iל־v_j שצמתי הביניים שלו לקוחים מתוך $\{v_1,\dots v_k\}$, כלומר $\delta_k(i,j)=\infty$.
נראה שמתקיים $D_k(i,j)=\infty$.
נניח בשלילה שלא, ולכן לפי הגדרת D מתקיים:
$D_k(i,j)$ סופי או $D_{k-1}(i,k)+D_{k-1}(k,j)$ סופי.
לפי הנחת האינדוקצייה, אם זה המצב אזי יש מסלול מv_i ל־v_j שצמתי הביניים שלו לא מתוך $\{v_1,\dots,v_k\}$ בסתירה להנחת השלילה.

דוגמה 4 Sequence Alignment

⋆ דוגמה: הוקלדה המחרוזת "ocurrance" ומוצע למשתמש התיקון ל־"ocurrence".

כיצד נתאים בין שתי המחרוזות ?

o	c	$\boxed{\text{-}}$	u	r	r	$\boxed{\text{a}}$	$\boxed{\text{-}}$	n	c	e
|	|	|	|	|	|	|	|	|	|	|
o	c	c	u	r	r	$\boxed{\text{e}}$	$\boxed{\text{e}}$	n	c	e
שלושה תוים שלא התאימו										
o	c	-	u	r	r	$\boxed{\text{a}}$	n	c	e	
|	|	|	|	|	|	|	|	|	|	
o	c	c	u	r	r	e	n	c	e	
תו אחד שלא התאים + התאמה שאינה בין אותו תו										

<u>נתון:</u> שתי מחרוזות $\begin{cases} x=x_1x_2\dots x_n \\ y=y_1y_2\dots y_m \end{cases}$

שידוד M בין $\{1,\dots,n\}$ ו־$\{1,\dots,m\}$ יהיה חוקי אם:

1. כל אינדקס מופיע לכל היותר פעם אחת ב־M.
2. $(i,j),(i',j')\in M$ ו־$i<i' \Leftarrow j<j'$.

שאלה:

כיצד נכמת ערך שידוד חוקי?
אילו יש מיקום שאינו מותאם, אזי נשלם עליו $\delta>0$.
$(i,j)\in M \Leftarrow$ קיימת טבלה שמציימת מהי עלות השידוד של $\alpha(x_i,y_j)$.
<u>מטרה:</u> למצוא שידוד חוקי זול ביותר בין x ל־y.
<u>הבחנה:</u> נתונות $x=x_1\dots,x_n$ ו־$y=y_1\dots y_m$ אזי:
$(n,m)\in M$ או שלפחות אחד מ־x_n ו־y_nלא משודד (הוכחה, כיוון שאין הצטלבויות).
תתי הבעיות שענייננו אותנו הן רשאיות(תאופיין ע"י i ו־j)

$$\begin{cases} x_1\dots x_i \\ y_1\dots y_j \end{cases}$$

נגדיר את $A(i,j)$ להיות עלות שידוך חוקי זול ביותר בין $x_1 \dots x_i$ ו־$y_1 \dots y_j$.

$$B(i,j) = \begin{cases} \delta \cdot j & i = 0 \\ \delta \cdot i & j = 0 \\ min\{\alpha(x_i, y_j) + B(i-1, j-1), \delta + B(i-1, j), \delta + B(i, j-1)\} & \text{otherwise} \end{cases}$$

לכל $1 \leq i \leq n$ ו־$1 \leq j \leq m$: $A(i,j) = B(i,j)$.

שאלה:

כיצד לחשב את B?

	1	2	3	$\cdots$	m
1	0	δ	2δ		$m\delta$
2	δ				
3	2δ				
$\vdots$					
n	$n\delta$				$\underbrace{?}_{\text{output}}$

קיימים מספר סדרים עבורם זמן הריצה הוא $O(n \cdot m)$.
סיבוכיות זכרון $O(min\{n,m\})$.