המשך אלגוריתמים חמדנים

תזכורת: אלגוריתם חמדן לבעיית הפעילויות (שכבר ראינו)

- נתחזק:
  - $I =$ הקטעים שכבר נבחרו (פתרון נוכחי)
  - $S =$ קטעים שהם לא ב- I ולא נחתכים עם אף קטע ב- I
- אתחול:
  - $S \leftarrow \{1,\ldots,n\}$
  - $I \leftarrow \varnothing$
- צעד:
  - כל עוד $S \neq \varnothing$:
    - נבחר $i \in S$ עם זמן סיום מינימלי
    - $I \leftarrow I \cup \{i\}$
    - נסיר מ- S את i וכל קטע שנחתך עם i
- סיום:
  - כאשר $S = \varnothing$ נחזיר את I

מימוש וזמן ריצה

מימוש:

- נמיין את הקלט בסדר עולה לפי זמן סיום.
- נשמור את התוצאה במערך:

$$\left(f_{i_1} \leq f_{i_2} \leq \cdots \leq f_{i_n}\right)$$

- נתחזק מספר $F =$ זמן סיום אחרון בפתרון הנוכחי.
- נעבור על המערך הממוין:
  - לכל $1 \leq j \leq n$, נביט בקטע הנוכחי i_j.
  - נבדוק אם $s_{i_j} \geq F$.
    - אם כן:
      - נוסיף את קטע i_j לפתרון.
      - נעדכן $F \leftarrow f_{i_j}$.
    - אם לא: נדלג עליו.

זמן ריצה:

- מיון: $O(n \log n)$
- מעבר על מהערך: $O(1)$ לכל קטע

- ○ $\Leftarrow$ בסה"כ $O(n)$
- סה"כ: $O(n\log n)$

טענה: בכל שלב בשני האלגוריתמים, הקטע הבא שנוסף הוא הראשון (לפי זמן סיום) שלא נחתך עם אף קטע בפתרון הנוכחי.

[[בהנחה שזה נכון, האלגוריתם הנ"ל שקול לזה שהראינו בשיעור הקודם, ולכן הוא נכון.]]

סכמה להוכחת נכונות של אלגוריתם חמדן

- טענה: הפתרון הנוכחי בכל שלב מוכל באיזשהו פתרון אופטימלי
  - צריך להוכיח שמהטענה נובעת נכונות האלגוריתם
- הוכחת הטענה:
  באינדוקציה על שלבי האלגוריתם.
  צעד אינדוקטיבי – טיעון החלפה:
  לקחת פתרון אופטימלי קודם, ואם צריך, לבצע שינוי מקומי כדי לקבל פתרון אופטימלי חדש שמכיל גם את האיבר שהוספנו.
  [[כלומר, מתבוננים בפתרון האופטימלי ובאיבר החדש – אם הוא כבר מכיל אותו, אז סיימנו; אחרת, מבצעים החלפה אם איבר אחר שנמצא בפתרון האופטימלי כך שעדיין יתקבל פתרון אופטימלי.]]

בעיית חלוקת הפעילויות

בעיה באופי דומה: במקום לנסות לשבץ כמה שיותר פעילויות באולם אחד, רוצים לשבץ את **כל** הפעילויות במספר אולמות, ולעשות זאת כך שמספר האולמות הדרושים יהיה מינימלי.

- מופע: סדרה של קטעים פתוחים $1,\dots,n$ הניתנים ע"י זמני התחלה וסיום (s_i,f_i).
- פתרון חוקי: חלוקה של הקלת לתתי-קבוצות $\{1,\dots,n\}\supseteq C_1,\dots,C_d$ (כלומר כך שלכל $i\neq j$ מתקיים $C_i\cap C_j=\varnothing$, וכן $\bigcup_{i=1}^{d}C_i=\{1,\dots,n\}$), ובכל קבוצה C_i, הקטעים הם זרים:

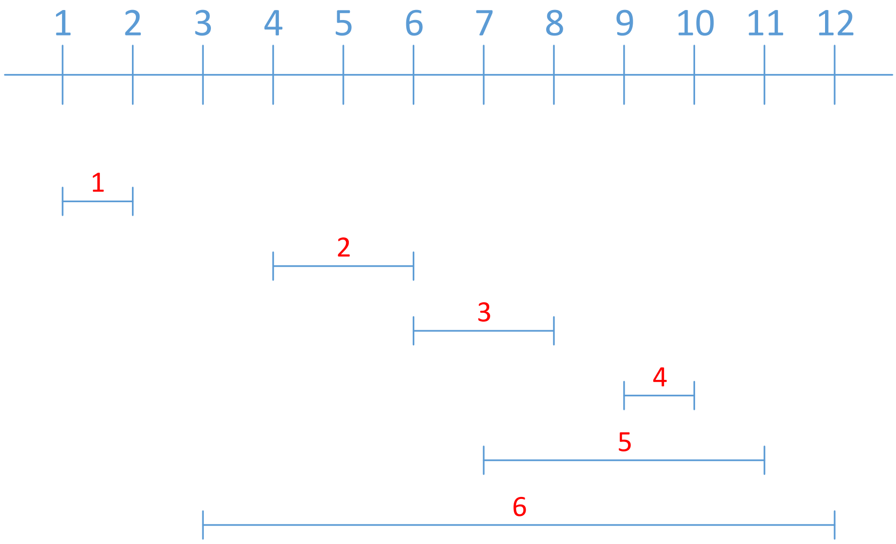
$$\forall j\neq k\in C_i:(s_j,f_j)\cap(s_k,f_k)=\varnothing$$

- יש למצוא: פתרון חוקי $C_1,\dots,C_d$ עם מס' מינימלי של קבוצות (כלומר d מינימלי).

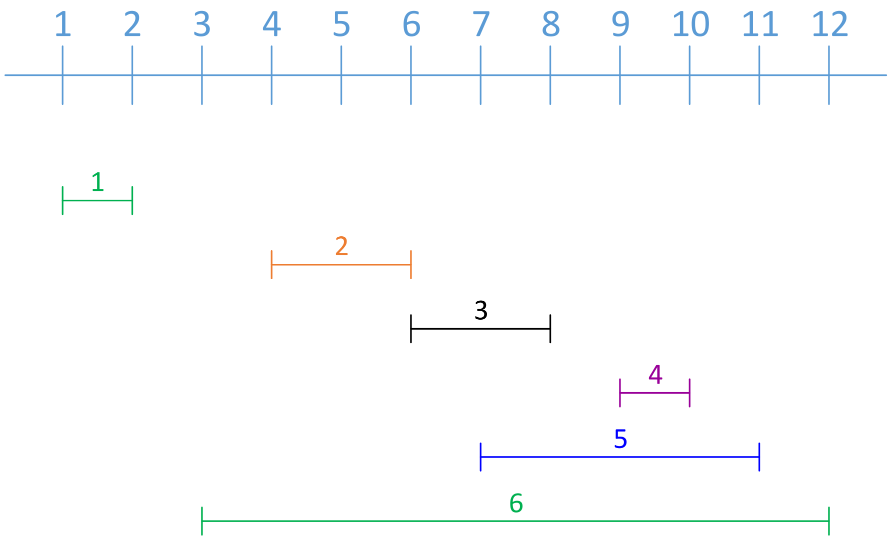
דוגמה

i	1	2	3	4	5	6
s_i	1	4	5	9	7	3
f_i	2	6	8	10	11	12

<u>ציור:</u>



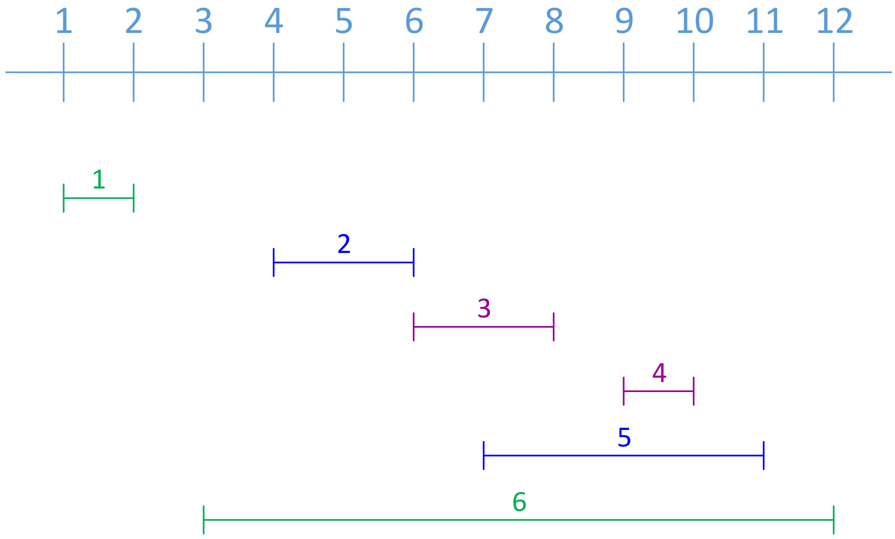
<u>פתרון חוקי:</u>



$C_1 = \{1,6\}$ $C_2 = \{2\}$ $C_3 = \{3\}$ $C_4 = \{4\}$ $C_5 = \{5\}$

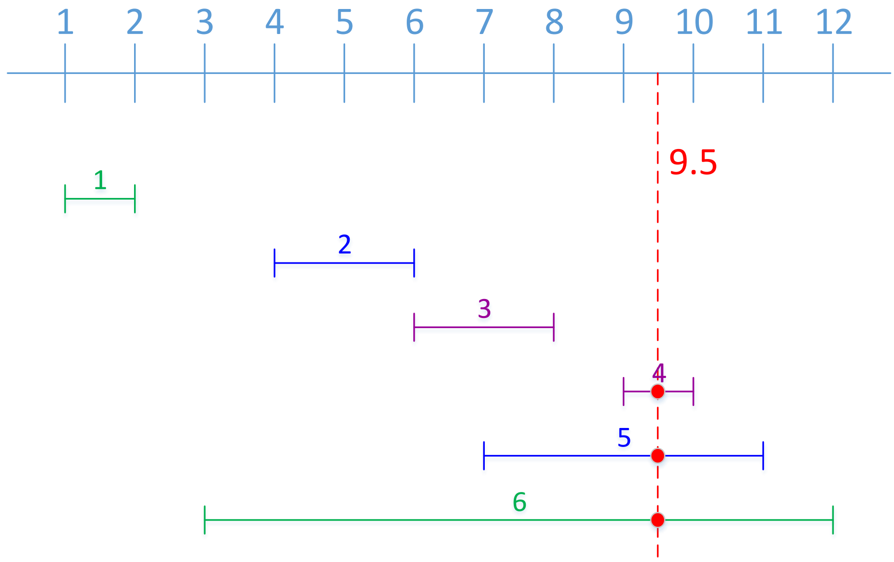
בפתרון זה יש 5 קבוצות.

<u>פתרון יותר טוב:</u>



$C_1 = \{1,6\}$ $C_2 = \{3,4\}$ $C_3 = \{2,5\}$

בפתרון זה יש 3 קבוצות, והוא גם אופטימלי.
ניתן לראות זאת כי ישנם 3 קטעים שכולם נחתכים, ולכן הם חייבים להיות בקבוצות שונות:



אלגוריתם חמדן לבעיה

- <u>נתחזק:</u>
  - $d=$ מס' הקבוצות בפתרון הנוכחי
  - $C_1,\dots,C_d$
- <u>אתחול:</u>
  - $d \leftarrow 0$
- <u>צעד:</u>
  - כל עוד יש פעילויות שלא שובצו:
    - נבחר מתוך הפעילויות הפנויות פעילות j <u>לפי כלל בחירה</u>...

- אם קיימת קבוצה C_i שניתן להוסיף לה את j (יתכן שיש כמה; אם כן, נבחר שרירותית אחת מהן) אזי:
  - $C_i \leftarrow C_i \cup \{j\}$
- אחרת:
  - $d \leftarrow d+1$
  - $C_d \leftarrow \{j\}$

- <u>סיום</u>:
  - נחזיר $C_1,\dots,C_d$

אפשרויות לכלל בחירה

1. <u>הפעילות שמסתיימת ראשונה [כפי שעשינו באלגוריתם הקודם]</u>:
נסה על הדוגמה שהראינו ונקבל:

$$C_1 = \{1,2,4\}$$
$$C_2 = \{3\}$$
$$C_3 = \{5\}$$
$$C_4 = \{6\}$$

כפי שראינו, בדוגמה זו ניתן למצוא פתרון עם 3 קבוצות, לכן כלל בחירה זה בהחלט לא בהכרח מניב פתרון אופטימלי.

2. <u>הפעילות **שמתחילה** ראשונה</u>:
למרות שפסלנו אותו באלגוריתם הקודם, הפעם כלל זה כן יעבוד.
למשל, בדוגמה שראינו, הקבוצות לאחר כל צעד של האלגוריתם עם כלל זה תהיינה:
   1. $C_1 = \{1\}$
   2. $C_1 = \{1,6\}$
   3. $C_2 = \{2\}$ $C_1 = \{1,6\}$
   4. $C_2 = \{2,3\}$ $C_1 = \{1,6\}$
   5. $C_3 = \{5\}$ $C_2 = \{2,3\}$ $C_1 = \{1,6\}$
   6. $C_3 = \{5\}$ $C_2 = \{2,3,4\}$ $C_1 = \{1,6\}$

קיבלנו פתרון שונה משראינו מקודם, אך הוא גם אחד הפתרונות האופטימליים האפשריים.

הגדרות ואבחנות

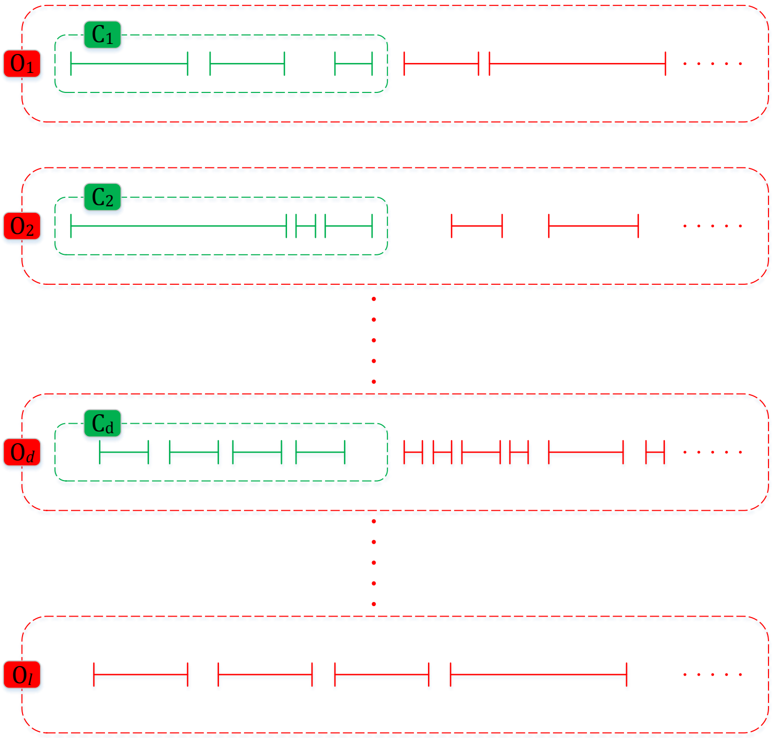
- <u>הגדרה</u>: עומק של קלט = מס' מקסימלי של קטעים שמתרחשים בו-זמנית.
- <u>אבחנה 1</u>: מס' הקבוצות בכל פתרון חוקי $\le$ עומק הקלט [לפי עקרון שובך היונים].
- <u>אבחנה 2</u>: האלגוריתם מחזיר פתרון חוקי [אנחנו לא כותבים לכך הוכחה פורמלית כי זה נובע די מידית מהגדרת האלגוריתם (הרי אנחנו שמים פעילות בקבוצה רק אם חוקי לעשות זאת, ומשבצים בדיוק כל פעילות בדיוק בקבוצה אחת, לכן התוצאה מהווה חלוקה)].

הוכחה לפי הסכימה

- <u>משפט</u>: האלגוריתם מחזיר פתרון אופטימלי
- <u>טענת עזר</u>: בכל שלב קיים פתרון אופטימלי $O=\{O_1,\dots,O_l\}$ כך שלכל C_i בפתרון הנוכחי מתקיים $C_i \subseteq O_i$.

[הסבר: זה לא נכון שבכל שלב הפתרון שלנו מוכל בפתרון אופטימלי כלשהו, כי למשל לאחר השלב הראשון יש בפתרון שלנו את הקבוצה $\{1\}$, ואין שום פתרון אופטימלי שבו פעילות מס' 1 נמצאת לבדה. לעומת זאת, מה שכן קורה, זה שבכל שלב, כל אחת מהקבוצות בפתרון שלנו מוכלת בקבוצות של פתרון אופטימלי כלשהו.]

הוכחת המשפט



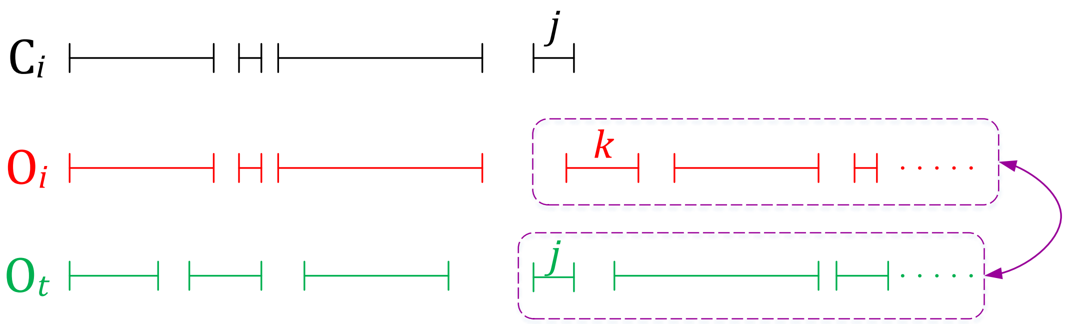
לפי טענת העזר, קיים פתרון אופטימלי O כך ש-$C_i \subseteq O_i$ (כאשר $C_1,\dots,C_d$ החלוקה שהאלגוריתם מחזיר). כיוון ש-C ו-O שתיהן חלוקות, $\forall i \;\; C_i = O_i$ [[כי אם קיים i כך ש-$O_i \setminus C_i \neq \varnothing$ אזי עבור $k \in O_i \setminus C_i$ קיים j כך ש-$k \in C_j$, ואז $k \in O_j$ ומקבלים כי $O_i \cap O_j \neq \varnothing$ ולכן $O_i = O_j$ (מכך ש-O היא חלוקה) ואז $C_i = C_j$ ולכן $k \in C_i$, בסתירה]].

הוכחת טענת העזר [בנפנוף ידיים]

[באינדוקציה.

בצעד האינדוקציה יש לנו C_i שמכיל את j אך O_i לא מכיל את j (ונניח מכיל k במקום, המתנגש עם j).

אז ניקח קבוצה O_t שכן מכילה את j (יש כזו כי O היא חלוקה של כל $\{1,\dots,n\}$), ונחליף את כל ה"זנב" של O_i החל מ- k עם כל הזנב של O_t החל מ- j .]



הוכחה נוספת (לא קשורה למה שהוכחנו עד כה) – כי אנחנו יכולים [במיוחד בשבילך, רוני ☺]

- <u>משפט</u>: האלגוריתם מחזיר פתרון אופטימלי.
- <u>טענת עזר</u>: האלגוריתם מוצא פתרון עם מס' קבוצת השווה לעומק הקלט.

<u>הוכחת המשפט</u>:
כל פתרון חייב להכיל לפחות מס' כזה של קבוצות (מאבחנה 1), ולכן האלגוריתם מחזיר פתרון אופטימלי.
[[כלומר, לפי טענת העזר, מס' הקבוצות שהאלגוריתם מחזיר שווה לעומק הקלט, לכן כיוון שמאבחנה מס' 1 זהו המספר המינימלי של קבוצות שיכולות להיות בפתרון חוקי, החלוקה שהאלגוריתם מחזיר היא בעלת מס' הקבוצות הקטן ביותר – כלומר היא אופטימלית.]]

<u>הוכחת טענת העזר</u>:
מאבחנה 1, מס' הקבוצות $d \geq$ עומק הקלט.
צ"ל כי מס' הקבוצות $d \leq$ עומק הקלט.

יהי j הקטע שבגללו פתחנו את C_d (הקבוצה האחרונה).
לפי האלגוריתם, C_d נפתחה כי כל קבוצה C_i ($i<d$) הכילה קטע k_i שנחתך עם j .
לפי סדר המעבר על הקטעים, $s_j \geq \max_{i<d}\{s_{k_i}\}$.
זמן ההתחלה של j

מכך שכל k_i נחתך עם j (וְ- k_i מתחיל לפני j) נובע ש- $f_{k_i} > s_j$.
לסיכום:

$$\max_{i<d}\{s_{k_i}\} \leq s_j < \min_{i<d}\{f_{k_i}\}$$

כלומר קיימת נקודה $(s_j+\varepsilon)$ שנמצאת בכל הקטעים $j,k_1,k_2,\dots,k_{d-1}$.
לכן עומק הקלט $\geq \left|\{j,k_1,\dots,k_{d-1}\}\right| = d$.

מ.ש.ל. ☐

]] הסבר:

- j נחתך עם $d-1$ הקטעים $k_1,\ldots,k_{d-1}$ (אחרת, אם לא היו כאלה קטעים בקבוצות C_i, לא היינו צריכים לפתוח את קבוצה C_d עבור קטע j).
- כל שני קטעים שנחתכים, בכלל, מכילים תת-קטע משותף, כלומר כל אחד מהם מתחיל לפני שהשני נגמר ונגמר אחרי שהשני מתחיל.
- בפרט, כל קטע k_i $(1\le i\le d-1)$ נגמר אחרי ש-j מתחיל – כלומר $s_j < f_{k_i}$ (שימו לב: זה $<$, ולא $\le$, כי אם $s_j = f_{k_i}$ אז הקטעים לא נחתכים). נסמן $f=\min\left\{\min\limits_{i<d}\left\{f_{k_i}\right\}, f_j\right\}$. אז $s_j < f$.
- בנוסף, מכך שתמיד בוחרים באלגוריתם את הקטע שמתחיל הכי מוקדם, כל הקטעים k_i, שנבחרו לפני קטע j, מתחילים לפניו – כלומר $s_{k_i} \le s_j$.
- לכן ניתן למצוא נקודה שנמצאת קצת אחרי s_j, שכל הקטעים נמצאים בה. למשל, הנקודה $\frac{s_j+f}{2}$ (שנמצאת בדיוק באמצע בין s_j לבין f) נמצאת בכל d הקטעים $\{j,k_1,\ldots,k_{d-1}\}$, ולכן עומק הקלט הוא לפחות כגודל קבוצה זו, שהוא d.

[[

מימוש וזמן ריצה

<u>מימוש:</u>

- <u>נתחזק</u> לכל קבוצה זוג: (C_i,F_i), כאשר:
  - $C_i=$ רשימה מקושרת
  - $F_i=$ זמן סיום אחרון של קטע ב-C_i

נשים לב שניתן להוסיף את j לקבוצה קיימת $\Leftrightarrow$ יש i כך ש-$F_i \le s_j$ $\Leftrightarrow$ $\min\limits_i F_i \le s_j$ (מינימום בין כל קבוצות i הקיימות).

- <u>נתחזק</u> ערימה Q עם זוגות (C_i,F_i), כאשר המפתחות לפיהם תעבוד הערימה יהיו ערכי F_i.
- <u>מיון</u>: נמיין את כל הקטעים לפי זמן התחלה ונשמור במערך.
- <u>אתחול</u>: $Q\leftarrow\varnothing$
- נעבור על המערך, ולכל קטע j:
  - $(C,F)\leftarrow \text{Extract\_Min}(Q)$
  - אם $F\le s_j$ אז:
    - $C\leftarrow C\cup\{j\}$
    - $F\leftarrow f_j$

- נכניס מחדש את (C,F) לערימה
- אחרת:
- ניצור זוג חדש $(\{j\}, f_j)$ ונכניסהו לערימה.

זמן ריצה:

- אתחול: מיון – $O(n\log n)$
- כל צעד: $O(\log n)$ לפעולות ערימה
  - לכן סה"כ $O(n\log n)$ [[כי הלולאה רצה n פעמים]]
- החזרת הפלט: $O(n)$
- סה"כ: $O(n\log n)$