קושי-NP וְ-Subset Sum

בעיה NP -שלמה היא בעיה "קשה כמו כל בעיה ב- NP ".

בעיות NP-שלמות שראינו:

- SAT
- 3SAT
- IS
- VC
- CLIQUE

ע"מ להראות שבעיה היא NP -קשה, מספיק להראות רדוקציה פולינומית מבעיה **כלשהי** שהיא NP -קשה אליה. לכן חשוב להכיר בעיות NP -קשות.

<u>חשוב מאוד לעשות את הרדוקציה בכיוון הנכון</u> – כלומר **מ**הבעיה שאנחנו יודעים שהיא NP -קשה **ל**בעיה שאנו רוצים להראות שהיא NP -קשה.

למשל, בשביל להראות ש- $SUSU$ [= Subset Sum] היא NP -קשה, אפשר להראות:

$$SAT \leq_p SUSU$$

[**מ-** SAT **ל-** $SUSU$.]

[[דרך לזכור את הכיוון: רוצים להראות ש- $SUSU$ קשה לפחות כמו SAT ; תחשבו על $\leq_p$ כיחס סדר של קושי הבעיה.]]

דרך פתרון כללית להראות ששפה היא NP-שלמה

נתון A . הוכח ש- $A \in NP-Comp$.

איך עושים את זה?

1) נוכיח $A \in NP$:
   - אלגוריתם אימות (פולינומי)
   - עֵד
   - פולינום לגודל העד
2) נראה $A \in NP-Hard$ ע"י בחירת $B \in NP-Hard$ והוכחת $B \leq_p A$.

[טעות נפוצה היא להוכיח רק את 2.
[ציור]
אם מוכיחים רק את 2, מקבלים שהבעיה נמצאת בקבוצה $NP-Hard$ (אשר מסומנת ב[צבע רנדומלי כלשהו שאני אבחר אם וכשאשרטט את זה] בציור).]

בעיית SUSU

נתונים T, $C=(c_1,\dots,c_n)$ ($c_i,T\in\mathbb{N}$).

צ"ל האם קיימת תת-סדרה של C שסכומה שווה ל- T.

[נראה בכיתה כי $SUSU\in NP-Comp$.]

בעיית Partition

נתון $A=(a_1,\dots,a_n)$ ($a_i\in\mathbb{N}$).

האם ניתן לחלק את A לשתי תתי-סדרות שסכומן שווה זה לזה?

צ"ל כי $Partition\in NP-Comp$.

1) $Partition\in NP$.
אלגוריתם אימות:
מקבל עד = שתי תתי-סדרות, יוודא שהן תקינות, יבדוק את סכומן שווה ויענה בהתאם.
ברור שזה פולינומי בגודל הקלט.

2) נוכיח $Partition\in NP-Hard$, כלומר לכל $A\in NP$ מתקיים $A\le_p Partition$.
מספיק להראות $SUSU\le_p Partition$.

[[נגדיר $S=$ סכום האיברים.
נניח כרגע כי $T>\frac{1}{2}S$.
[ציור]
אז חסר לנו איברים בגודל $2T-S$ ע"מ שנקבל שתי עמודות שוות.
לכן נוסיף איבר אחד בעל ערך זה.
נניח שקיבלנו פתרון לבעיה החדשה. למה יש פתרון לבעיה המקורית?
כי במקרה זה יש לנו שתי קבוצות של איברים בקבוצה החדשה שהסכום של כל אחת מהן הוא T. אמנם הוספנו איבר חדש, אבל הוא יכול להיות רק באחת מהן, לכן בקבוצה השנייה יש איברים שכולם מהקבוצה המקורית, וסכומם הוא T – כלומר יש פתרון לבעיה המקורית.]]

נחשב $S=\sum_{i=1}^{n}c_i$.

נפלוט $A=C\cup(t)$ [[כלומר שרשור של האיבר t לסוף הסדרה A]] כאשר:

- אם $T>\frac{1}{2}S$ אז $t=2T-S$
- אחרת $t=S-2T$

נוכיח $x\in SUSU\Leftrightarrow f(x)\in Partition$.

- <u>$x\in SUSU\Rightarrow f(x)\in Partition$</u>:

עבור המקרה בו $T < \frac{1}{2}S$:

אם $x \in SUSU$ אזי קיימת קבוצת אינדקסים $D = \{i_1, \ldots, i_k\}$, $i_j \neq i_r$ לכל $j \neq r$ [[שזה בכל מקרה קורה, כי זו קבוצה, אז אין בה את אותו איבר פעמיים...]], כך ש:

$$\sum_{j \in D} c_j = T < \frac{1}{2}S = \frac{1}{2}\sum_{c_j \in C} c_j$$

נבחר:

$$D_1 = (c_i : i \in D) \cup (t)$$
$$D_2 = (c_i : i \notin D)$$

[[כלומר בוחרים תת-סדרה שמכילה את האיברים עם אינדקסים ב-D שלסופה משרשרים את האיבר החדש t, ועוד תת-סדרה שמכילה את שאר איברי הסדרה המקורית (הכתיב כאן לא הכי פורמלי).]]

ברור כי D_1, D_2 זרות [[מבחינת האיברים מהקבוצה שנמצאים שם, לא מבחינת ערכי האיברים (יכולים להיות בסדרה המקורית ערכים שמופיעים כמה פעמים)]].

$$\sum_{c_i \in D_1} c_i = T + t = T + S - 2T = S - T$$
$$\sum_{c_i \in D_2} c_i = S - T$$

[עבור המקרה בו $T > \frac{1}{2}S$ ההוכחה דומה, ואת המקרה בו $T = \frac{1}{2}S$ ניתן להוסיף לאחד מהשניים האחרים וזה עדיין יעבוד – אבל למעשה אם $T = \frac{1}{2}S$, הבעיה הנתונה שקולה לבעיית $Partition$ ממילא, לכן הפתרונות שקולים.]

- <u>$x \in SUSU \Leftarrow f(x) \in Partition$:</u>

[ניתן להוכיח גם $x \notin SUSU \Rightarrow f(x) \notin Partition$ – אבל שימו לב לא להוכיח $x \notin SUSU \Leftarrow f(x) \notin Partition$, כי זה בכלל המקרה הקודם.]

שוב עבור המקרה בו $T < \frac{1}{2}S$:

נניח כי $f(x) \in Partition$. לכן קיימת חלוקה של A לשתי סדרות, D_1, D_2 כך ש-D_1, D_2 זרות וכן:

$$\sum_{d\in D_1} d = \sum_{d\in D_2} d = \frac{1}{2}\sum_{a\in A} a$$

נניח בה"כ כי $t\in D_1$ ונבחר:

$$D=\{i : c_i \in D_1\}$$

איברי D זרים (מחוקיות D_1 [[כלומר מכך שהאיברים ב- D_1 זרים זה לזה]]).
סכום איברי D_1 הוא:

$$\sum_{a_j\in D_1} a_j$$

וְ- $t=S-2T$ [במקרה שלנו], לכן:

$$\sum_{j\in D} a_j = \sum_{a_j\in D_1} a_j - t = S-T-t = S-T-(S-2T) = T$$

[[כלומר קיבלנו שיש תת-סדרה של C שסכום איבריה הוא T, לכן $x\in SUSU$.]]

[נשים לב שכיוון ש- $SUSU$ וְ- $Partition$ שתיהן NP-שלמות, יש גם את הרדוקציה ההפוכה (זה נכון לכל שתי שפות ב- $NP-Comp$).]

בעיית Knapsack ("התרמיל ולא הגנב" ☺)

נתון $W,P,w,p,A=(a_1,\ldots,a_k)$:

- w – פונקציית משקל
- p – פונקציית ערך
- W – משקל מקסימלי
- P – ערך מינימלי

צ"ל: האם קיימת קבוצת אינדקסים $B=\{i_1,\ldots,i_k\}$ (אינדקסים זרים) כך ש:

$$\sum_{j\in B} w(a_j) \le W$$

$$\sum_{j\in B} w(a_j) \ge P$$

הוכיחו כי $Knapsack \in NP-Comp$.

<u>פתרון:</u>

שוב, להראות ש- $Knapsack \in NP$ זה פשוט [אבל תזכרו **שצריך לעשות זאת!**].

נראה:

$$SUSU \leq_p Knapsack$$

בהינתן מופע ל- $SUSU$, T, $C=(c_1,\dots,c_n)$, נגדיר $W=P=T$, ונייצר $A=(a_1,\dots,a_n)$ כאשר:

$$p(a_i)=w(a_i)=c_i$$

הרדוקציה פולינומית [כי בניה זו ניתן לבצע בזמן ליניארי בגודל הקלט].

- <u>$f(x)\in KS \Leftarrow x\in SUSU$:</u>

אם $x\in SUSU$ אזי קיימת קבוצת אינדקסים $B=\{i_1,\dots,i_k\}$ כך ש:

$$\sum_{i\in B} c_i = T$$

מהגדרה:

$$\sum_{i\in B} w(a_i)=\sum_{i\in B} c_i = T$$
$$\sum_{i\in B} p(a_i)=\sum_{i\in B} c_i = T$$

וקיבלנו פתרון חוקי ל- KS.

- <u>$x\in SUSU \Leftarrow f(x)\in KS$:</u>

במקרה זה קיימת קבוצת אינדקסים $B=\{i_1,\dots,i_k\}$ כך ש:

$$\sum_{i\in B} w(a_i)\leq W$$
$$\sum_{i\in B} p(a_i)\geq P$$

נגדיר $D=B$ ונקבל:

$$\sum_{j\in D} c_j = \sum_{j\in B} w(a_j)\leq W = T$$
$$\sum_{j\in D} c_j = \sum_{j\in B} p(a_j)\geq P = T$$

לכן [[מאנטי-סימטריה של יחס אי-השוויון בטבעיים]]:

$$\sum_{j\in D} c_j = T$$

שאלה [האם Knapsack פולינומי?]

כשעשינו פתרון לבעיית $Knapsack$ בתכנון דינאמי, קיבלנו משהו שנראה פולינומי – זמן הריצה היה:

$$O(n \cdot W)$$

[[זה הגיע ממטריצה בגודל $n \times W$]].

האם זה פולינומי בגודל הקלט?
השאלה שיש לענות עליה לשם לקבת תשובה לכך היא: **מהו גודל הקלט?**

גודל הקלט הוא:

$$O(n + \log W + \log P)$$

(כי מס' W מיוצג ע"י $\log W$ ביטים).

נזכור שמתקיים:

$$2^{\log W} = W$$

לכן אם $\log W \geq n$, הפתרון שהצגנו בהחלט **לא פולינומי בגודל הקלט** (אלא אקספוננציאלי).

השמה מספקת-לא מספקת

שפת $NAE - SAT$ מוגדרת בתור שפת כל הנוסחאות שיש הצבה חוקית שמספקת אותן, בה בכל פסוקית יש לפחות ליטרל אחד המקבל ערך $\mathcal{F}$.

לדוגמה עבור הנוסחה:

$$(x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_1 \vee \neg x_2 \vee x_3)$$

ההצבה:

$$x_1 = x_2 = x_3 = \mathcal{T}$$

הינה מספקת-לא מספקת, שכן היא גורמת לכל הפסוקיות לקבל ערך $\mathcal{T}$ אבל קיים לפחות ליטרל אחד בכל אחת שמקבל ערך $\mathcal{F}$:

$$(\mathcal{T} \vee \mathcal{T} \vee \mathcal{F}) \wedge (\mathcal{F} \vee \mathcal{F} \vee \mathcal{T})$$

השאלה: האם $NAE - SAT \in NP - Comp$?

[זכרו: יש גם להראות כי $NAE - SAT \in NP$ לשם כך, למרות שכאן אנו נתמקד בלהראות כי $NAE - SAT \in NP - Hard$.]

נראה רדוקציה:

$$3SAT \leq_p NAE - 3SAT$$

[$NAE - 3SAT$ היא אותה בעיה רק מוגבלת לפסוקיות בנות 3 ליטרלים.]

[הרעיון: עבור פסוקית i, אותה נסמן ע"י $\left(\varphi_{i,1} \vee \varphi_{i,2} \vee \varphi_{i,3}\right)$, נבצע את ההמרה הבאה:

$$\left(\varphi_{i,1} \vee \varphi_{i,2} \vee \varphi_{i,3}\right) \xrightarrow{\text{Reduction}} \left(\varphi_{i,1} \vee \varphi_{i,2} \vee t_i\right) \wedge \left(\overline{t_i} \vee \varphi_{i,2} \vee b\right)$$

ונראה שזה עובד.]

בהינתן נוסחה $\varphi = c_1 \wedge c_2 \wedge \cdots \wedge c_m$ נסמן $c_i = \left(\varphi_{i,1} \vee \varphi_{i,2} \vee \varphi_{i,3}\right)$.

תהי $\varphi \in 3SAT$.
אזי קיימת הצבה המספקת את φ.
[[כאן הלוח נמחק במסתוריות; אנא פנו למסמך התרגול כדי לראות את ההוכחה.]]

אם $f(\varphi) \in NAE - 3SAT$:

- אם $b = \mathcal{F}$, הצבת ערכי המשתנים מ-φ' ב-φ תייצר הצבה מספקת.
- אם $b = \mathcal{T}$, יתכן שאותה הצבה לא תספק את φ – אבל!

אם **הופכים** את כל ערכי המשתנים בהצבה, מקבלים הצבה **שכן** מספקת את φ, וגם זו הצבה מספקת-לא מספקת ל-φ', שכן ההצבה ההפוכה להצבה מספקת-לא מספקת גם היא מספקת-לא מספקת.