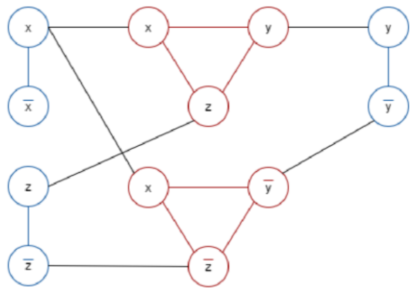
הרצאה 8 – יום ראשון 17.8

$VC \in NPC$

בעיית *vertex-cover* (ראינו באלגו 2): תת-קבוצה של קודקודים כך שכל צלע נוגעת בלפחות קודקוד אחד בקבוצה.

$$VC \coloneqq \{\langle G, k\rangle \mid G \text{ has a vx cover of size } k\}$$

כדי להראות שהשפה NPH, נעשה רדוקציה משפת $3\text{-}SAT$. לכל משתנה x נייצר גאדג'ט: שני קודקודים, $x, \bar{x}$, עם צלע ביניהם (נקרא להן צלעות של המשתנים). לכל פסוקית נייצר גאדג'ט – משולש: קודקודים לפי הליטרלים, ונחבר ביניהם צלעות (צלעות של הפסוקיות). ונחבר צלעות בין כל קודקוד של ליטרל (מהגאדג'ט של המשתנים) לקודקודים המתאימים בגאדג'ט של הפסוקיות (צלעות מחברות). לדוגמה עבור $\varphi = (x \lor y \lor z) \land (x \lor \bar{y} \lor \bar{z})$:



הוכחת נכונות: ההשמה המספקת תתורגם לקודקודים שלוקחים לכיסוי. מהגאדג'טים של המשתנים, ניקח את הקודקודים שקיבלו T. אז כל הצלעות של הגאדג'טים של המשתנים מכוסות. מהגאדג'טים של הפסוקיות, נבחר את אחד הקודקודים שקיבלו T ואותו **לא ניקח**, וניקח את השניים האחרים. בגלל שיש לפחות ליטרל אחד מסופק בכל פסוקית, אז בכל פסוקית יש לכל היותר שניים לא מסופקים. ואנחנו לא לוקחים אחד שמסופק. אז כל קודקוד של ליטרל שלא מסופק, בהכרח ייבחר. בסה"כ ניקח $2n + m$ קודקודים, וזה יהיה ה-k שלנו.

אם φ ספיקה, אז לקחנו קודקוד אחד מכל גאדג'ט של משתנה – זה מכסה את הצלעות של המשתנים. ומכל משולש, לקחנו שני קודקודים – זה מכסה את הצלעות של הפסוקיות. מה לגבי הצלעות המחברות? אם קודקוד של משולש מתאים לקודקוד של המשתנה שבחרנו (כלומר הוא קיבל T תחת ההשמה) אז הצלע ביניהם מכוסה בגלל הקודקוד של המשתנה. אם קודקוד של משולש הוא הנגדי (קיבל F), אז הוא יהיה אחד מבין השניים שבחרנו, וזה מכסה את הצלע.

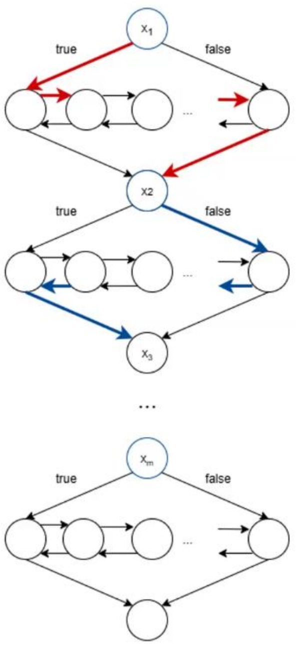
אם הכיסוי המינימלי ב-G הוא בגודל לכל היותר $2n + m$: ראשית, נשים לב שכל משולש דורש לפחות שני קודקודים, אז בחרנו לפחות $2n$ קודקודים. כלומר אפשר לקחת לכל היותר m קודקודים אחרים מתוך הקודקודים של המשתנים. מצד שני, חייבים תמיד לקחת לפחות אחד מתוך כל זוג של קודקודים של המשתנים, ויש m כאלה. אז מכל זוג בחרנו בדיוק אחד. זה אומר לנו אם המשתנה T או F. אז לקחנו בדיוק m קודקודים מתוך הזוגות האלה. זה אומר שאפשר לקחת לכל היותר עוד $2n$ קודקודים מהמשולשים. וניזכר שאמרנו שלקחנו לפחות $2n$, כלומר במשולשים בחרנו בדיוק $2n$.

אז בכל משולש, בדיוק קודקוד אחד לא נבחר לכיסוי. אם הוא יהיה F תחת ההשמה, אז גם הקודקוד שמתאים לו לא נבחר, ואז הצלע המחברת ביניהם לא מכוסה. כלומר, כל קודקוד שלא נבחר לכיסוי חייב לקבל T. אז לכל פסוקית יש ליטרל מסופק.

עבור פסוקית עם n משתנים ו-m פסוקיות, נייצר $2n + 3m$ קודקודים. אז גודל הגרף וזמן הייצור הם פולינומים בגודל הפסוקית.

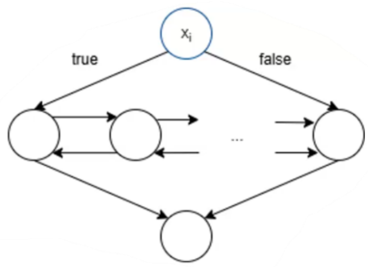
והשפה ב-NP – העד הוא קבוצת קודקודים, והבדיקה האם היא מכסה היא פולינומית. בסה"כ, $VC \in NPC$.

$HAMPATH \in NPC$

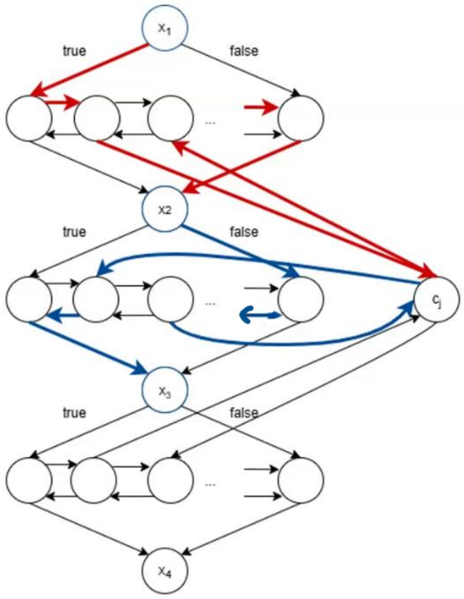


כדי להראות שהשפה NPH, נעשה רדוקציה מ- $3\text{-}SAT$. בהינתן פסוק $\varphi = c_1 \land c_2 \land \ldots \land c_n$ עם m משתנים, נבנה גרף.

לכל משתנה x_i נבנה תת-גרף:

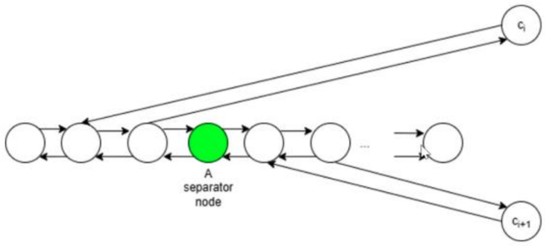


ברעיון, אם המשתנה קיבל T, אז המסלול ההמילטוני יעבור דרך הצלע *true*. ונחבר בין כל תתי הגרפים של המשתנים:



עוד לא התייחסנו למבנה של הנוסחה עצמה. לכל פסוקית, נוסיף קודקוד. אם x_1 מופיע בתור ליטרל חיובי ב- c_j, אז תהיה צלע מהגאדג'ט של x_1 ל-c_j ובחזרה מ- c_j ל- x_1 (לקודקוד הבא), בכיוון לפי המופע (חיובי או שלילי). כלומר, אם x_1 מופיע בתור ליטרל חיובי ב- c_j, אז הצלעות יהיו כך שאפשר לעבור דרך c_j רק אם המסלול עובר מהצלע החיובית של x_1. אם הוא היה מופיע בתור ליטרל שלילי, אז הצלעות היו כך שאפשר לעבור דרך c_j במסלול רק אם המסלול עובר מהצד השלילי של x_1. כלומר המסלול עובר דרך פסוקית c_j רק אם היא מסופקת. לדוגמה, עבור הפסוקית $c_j = (x_1 \vee \overline{x_2} \vee x_3)$:

כדי לוודא שהמסלול הוא רק בכיוון אחד ועובר דרך הקודקודים של הפסוקיות, בין כל זוג של קודקודים שמחובר לפסוקית נוסיף קודקוד מפריד:



זה גם קובע לנו כמה קודקודים יהיו בשורה הזאת באמצע – כל פסוקית "מקשרת" בין שני קודקודים ייחודיים, ויש אחת מפרידה. $O(3n)$. זה, כפול m משתנים – $O(3mn)$. ועוד קודקוד לכל פסוקית – m. סה"כ $O(3mn + m) = O(3mn)$. ומספר המשתנים פולינומי במספר הפסוקיות.

נכונות: אם הנוסחה ספיקה, אז לכל משתנה נלך מהצד לפי ההשמה, ולכל פסוקית נגיע דרך המשתנים שמסופקים אצלה, וזה מסלול המילטוני. אם הנוסחה לא ספיקה, אז יש פסוקית לא מסופקת – כלומר אי אפשר להגיע לפסוקית הזאת בלי לחזור על קודקודים.

וכמובן שהשפה ב-NP – בהינתן גרף, המסלול הוא העד והבדיקה פולינומית. בסה"כ, $HAMPATH$ היא NPC.

$\boldsymbol{SUBSET\text{-}SUM \in NPC}$

$$SUBSET\text{-}SUM \coloneqq \{\langle S, t\rangle \mid S \subseteq \mathbb{N}, \quad \exists S' \subseteq S : \Sigma S' = t\}$$

נעשה רדוקציה מ- $3\text{-}SAT$. בהינתן נוסחת $3\text{-}CNF$ עם m משתנים ו-n פסוקיות, נבנה מספרים כך שלפי "עמודות" (לפי המיקומים של הספרות) הדרך היחידה להגיע ל-t היא לקחת רק אחד מתוך הערכים T, F של כל איבר, וגם שלכל פסוקית, לפחות ליטרל אחד מסופק.

יש עמודה לכל משתנה ולכל פסוקית. לדוגמה עבור $\varphi = (x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \neg x_2 \vee \neg x_3)$, כל מספר יהיה מהצורה: a_1, a_2, a_3, b_1, b_2. שלוש הספרות השמאליות מייצגות את המשתנים, שתי הספרות הימניות מייצגות את הפסוקיות. לכל משתנה x_i נייצר שני מספרים:

$$y_i \coloneqq 10^{n+i-1} + \sum_{j : x_i \in c_j} 10^{j-1}, \qquad z_i \coloneqq 10^{n+i-1} + \sum_{j : \neg x_i \in c_j} 10^{j-1}$$

- y_i, שיש לו ספרה 1 במיקום של x_i ובמיקום של כל פסוקית שבה מופיע x_i, בתור ליטרל חיובי.
- z_i, שיש לו ספרה 1 במיקום של x_i ובמיקום של כל פסוקית שבה מופיע $\neg x_i$, בתור ליטרל שלילי.

לדוגמה עבור הנוסחה φ, נייצר:

$$y_1 \coloneqq 10011, \qquad y_2 \coloneqq 01010, \qquad y_3 \coloneqq 00110, \qquad z_1 \coloneqq 10000, \qquad z_2 \coloneqq 01001, \qquad z_3 \coloneqq 00101$$

בדיקת שפיות: רק אחד מתוך ה-m ספרות הראשונות אמור להיות 1, כל השאר צריכים להיות 0. ולכל i, ה-n ספרות האחרונות של y_i, z_i צריכות להיות משלימים אחד של השני. **המספרים עשרוניים**, לא בינאריים.

בנוסף, נגדיר $t \coloneqq (1)^m(3)^n$ (שרשור, לא חזקה). בדוגמה שלנו, $t \coloneqq 11133$.

עכשיו, הדרך היחידה לקבל סכום t זה לקחת רק אחד מתוך y_i, z_i – כלומר לבחור השמה T או F למשתנה, בלי סתירות. הבעיה היא, שלא צריך לבחור לכל פסוקית 3 ליטרלים מסופקים – מספיק רק אחד. אבל אם נגדיר את $t \coloneqq 11122$, אז זה לא מאפשר מצב שכל הליטרלים מסופקים. פיתרון: לכל פסוקית c_j נוסיף שני מספרים: (הערה – טכנית זה מולטי-קבוצה, כי בקבוצה אין כפילויות של אותו מספר).

$$g_j = h_j \coloneqq 10^{j-1}$$

בדוגמה שלנו, $g_1 = h_1 \coloneqq 01, g_2 = h_2 \coloneqq 10$. כלומר נוסיף $\{01,01,10,10\}$. זה מאפשר לנו "להשלים" את הסכום ל-t, גם אם יש פסוקית שלא כל הליטרלים שלה מסופקים. ובגלל שלמספרים האלה יש אחדות רק במקומות של הפסוקיות, זה לא מבטל את הדרישה שכל משתנה יקבל השמה תקינה.

לבסוף, נגדיר:

$$S \coloneqq \bigcup_{x_i} \{y_i, z_i\} \cup \bigcup_{c_j} \{g_j, h_j\}$$

בסה"כ, השמה מספקת מקבילה בדיוק לתת-מולטי-קבוצה בסכום t.

אם לא רוצים להשתמש במולטי-קבוצה, אפשר להגדיר את $t \coloneqq (1)^m(4)^n$, נגיד במקרה שלנו $t \coloneqq 11144$. ונגדיר: $h_j \coloneqq 20^{j-1}$, $g_j \coloneqq 10^{j-1}$.
אם ליטרל אחד מסופק, ניקח את g, h כדי להשלים ל-4. אם שני ליטרלים מסופקים, ניקח רק את h. אם כל הליטרלים מסופקים, ניקח רק את g. ואם אין ליטרלים מסופקים, אז $g + h$ לא ישלימו ל-4.

למה אנחנו מתייחסים למחלקות הזמן בתור מחלקות של שפות ב-R?

הרעיון הוא שאנחנו מגדירים שייכות למחלקת זמן לפי קיום מ"ט שמקבלת את השפה בזמן שמוגבל ע"י פונקציה כלשהי. כלומר המ"ט מגיעה למצב מקבל תוך מספר חסום של צעדים. אז נוכל לחסום את הריצה של המכונה באופן כללי – אם לא הגענו למצב מקבל תוך זמן מסויים (או מספר מסויים של צעדים), נדחה. כלומר תמיד נקבל תשובה, אז השפה שייכת ל-R. זה נקרא *stopwatch* (שעון עצר) של השפה. עבור מחלקות *P, NP, coNP* זה שעון עצר פולינומי.

מחלקות *EXPTIME, NEXPTIME*

מחלקת כל השפות שיש *DTM* שמכריע אותן, בזמן אקספוננציאלי בגודל הקלט:

$$\text{EXPTIME} \coloneqq \bigcup_{c=0}^{\infty} \text{DTIME}\left(O\left(2^{n^c}\right)\right)$$

מתקיים: $\text{NP} \subseteq \text{EXPTIME}$, כי סמלוץ של *NTM* ע"י *DTM* דורש זמן אקספוננציאלי.

באופן דומה, $\text{coNP} \subseteq \text{EXPTIME}$, כי כל שפה ב-*coNP* אפשר להכריע ע"י מעבר *brute-force* על כל העדים עבור ה-*NTM* של השפה המשלימה.

כלומר: אם שפה L ב-*coNP*, זה אומר שהשפה המשלימה שלה ב-*NP*. כלומר עבור ה-*NTM* של השפה המשלימה, לכל מילה בשפה יש עד. אבל לכל מילה ב-L, כל עד נכשל על ה-*NTM*. אז יש *DTM* אקספוננציאלי עבור L – לכל מילה, נעבור על כל העדים האפשריים (מספר אקספוננציאלי באורך הקלט) ונראה אם יש אחד שעובד.

מחלקת כל השפות שיש *NTM* שמכריע אותן, בזמן אקספוננציאלי בגודל הקלט:

$$\text{NEXPTIME} \coloneqq \bigcup_{c=0}^{\infty} \text{NTIME}\left(O\left(2^{n^c}\right)\right)$$

טענה: אם $\text{EXPTIME} \neq \text{NEXPTIME}$, אז $\text{P} \neq \text{NP}$

<u>הוכחה</u> – טענת הריפוד: נניח ש- $\text{EXPTIME} \neq \text{EXPTIME}$, ונב"ש ש- $\text{P} = \text{NP}$. תהי $L \in \text{NEXPTIME}$, כלומר קיים *NTM* כלשהו M שמכריע אותה בזמן 2^{n^c} עבור קבוע c כלשהו. נגדיר את השפה:

$$L_{\text{pad}} \coloneqq \left\{\langle x, 1^{2^{|x|^c}} \rangle \mid x \in L\right\}$$

כלומר, נשרשר $2^{|x|^c}$ אחדות אחרי x. ונגדיר *NTM* שמכריעה את השפה – נקרא לה M':

בהינתן קלט w, נבדוק אם w במבנה הנכון – האם יש $2^{|x|^c}$ אחדות. אם כן, נריץ את $M(x)$ ונחזיר את מה שהיא מחזירה. כלומר, הריצה של $M(x)$ (שהייתה מעריכית בגודל של x) היא פולינומית בגודל של המילה המרופדת. אז $L_{\text{pad}} \in \text{NP}$. ומההנחה ש- $\text{P} = \text{NP}$, נקבל ש- $L_{\text{pad}} \in \text{P}$.

אם $L_{\text{pad}} \in \text{P}$, אז יש *N DTM* שמכריעה אותה בזמן פולינומי. עכשיו, נוכל לבנות *DTM* N' עבור L: בהינתן x, נבנה את $w \coloneqq \langle x, 1^{2^{|x|^c}} \rangle$ ונריץ ונחזיר את $N(w)$. N' רצה בזמן $O\left(2^{|x|^c}\right)$ כלומר $L \in \text{EXPTIME}$. זה מתקיים עבור כל $L \in \text{NEXPTIME}$, כלומר $\text{NEXPTIME} \subseteq \text{EXPTIME}$. וכמובן ש- $\text{EXPTIME} \subseteq \text{NEXPTIME}$, כלומר בסה"כ $\text{EXPTIME} = \text{EXPTIME}$, סתירה.

משפט היררכיית הזמן – *The Time Hierarchy Theorem*

פונקציה $f \colon \mathbb{N} \to \mathbb{N}$ תיקרא **חשיבות בזמן** (*time constructable*) אם קיים *DTM* שבהינתן קלט $(1)^n$ (שרשור) עוצר אחרי בדיוק $f(n)$ צעדים (או לחלופין, כותבת את $f(x)$ על הסרט ועוצרת). אפשר לחשוב על f בתור פונקציה "סבירה" – מונוטונית שעולה עם n.

<u>משפט היררכיית הזמן</u>: לכל שתי פונקציות $f(n), g(n)$ שהן *time constructable* כך ש- $f(n) \cdot \log n = o\big(g(n)\big)$, מתקיים:

$$\mathrm{DTIME}\big(f(n)\big) \subsetneq \mathrm{DTIME}\big(g(n)\big)$$

(הסימון הוא "תת-קבוצה לא שווה", כלומר $\mathrm{DTIME}\big(f(n)\big) \subset \mathrm{DTIME}\big(g(n)\big)$ וגם $\mathrm{DTIME}\big(f(n)\big) \neq \mathrm{DTIME}\big(g(n)\big)$).

כלומר, קיימת שפה $L \in \mathrm{DTIME}\big(g(n)\big)$ כך ש- $L \notin \mathrm{DTIME}\big(f(n)\big)$. ככל שיש יותר זמן, יש יותר שפות שאפשר להכריע.

נכתוב את הטענה בצורה נוחה יותר: $\mathrm{DTIME}\Big(o\big(f(n)\log f(n)\big)\Big) \subsetneq \mathrm{DTIME}\big(f(n)\big)$.

אנחנו נרצה, בהינתן פונקציה $f(n)$, לבנות שפה שניתנת להכרעה בזמן $f(n)$ אבל לא בזמן $o\big(f(n)\log f(n)\big)$.

נוכיח בשני שלבים: בשלב הראשון נתמקד ברעיון ההוכחה, ונתעלם מה- $\log n$. בשלב השני נפרמל את ההוכחה.

<u>שלב א – שיטת האלכסון</u>: יהי *DTM* כלשהו N. נבנה *DTM*, M: בהינתן קלט $\langle N\rangle$ (כאשר $n \coloneqq |\langle N\rangle|$):

1. נריץ את $N(\langle N\rangle)$ למשך $f(n)$ צעדים. אם $N(\langle N\rangle)$ עוד לא עצר, נדחה.
2. אם N עצר ודחה, נקבל. אם N עצר וקיבל, נדחה.

מתוך הבנייה, $L(M) \in \mathrm{DTIME}\big(f(n)\big)$. כי הוא רץ רק עד $f(n)$ צעדים.

נב"ש ש- $L(M) \in \mathrm{DTIME}\Big(o\big(f(n)\log f(n)\big)\Big)$. כלומר קיים *DTM* M' שמכריע את L בזמן $o\big(f(n)\log f(n)\big)$.

ניזכר מה המשמעות של הסימון o – זה אומר שהחל מאורך מסויים של קלט, מתקיים $\frac{t(M)}{t(M')} \to \infty$. כלומר M' רצה הרבה יותר מהר מאשר M.

נניח שאורך הקידוד של $\langle M'\rangle$ הוא לפחות האורך הזה, ונריץ את $M(\langle M'\rangle)$. הזמן שלו זה $f(n)$, והפלט הוא $M(\langle M'\rangle) \neq M'(\langle M'\rangle)$ - שזו סתירה, כי בהגדרה יש להם את אותה שפה. כלומר ההנחה שגויה, אז $L(M) \notin \mathrm{DTIME}\Big(o\big(f(n)\log f(n)\big)\Big)$.

אמרנו שהזמן של M הוא $f(n)$, אז למה צריך את ה- $\log n$? כשאמרנו שמריצים את $N(\langle N\rangle)$ למשך $f(n)$ צעדים, צריך איכשהו לעקוב אחרי הצעדים האלו. הספירה הזו דורשת זמן $\log f(n)$. נעדכן את ההגדרה של M: בהינתן קלט $\langle N\rangle$ (כאשר $n \coloneqq |\langle N\rangle|$):

1. נחשב ונשמור את $g \coloneqq g(n) = f(n)\ /\ \log f(n)$, בתור קבוע.
2. נריץ את $N(\langle N\rangle)$ למשך g צעדים, ובכל צעד נחסיר 1 מ- g. אם $N(\langle N\rangle)$ עוד לא עצר, נדחה.
3. אם N עצר ודחה, נקבל. אם N עצר וקיבל, נדחה.

אז $L(M) \in \mathrm{DTIME}\big(f(n)\log f(n)\ /\ \log f(n)\big) = \mathrm{DTIME}\big(f(n)\big)$. נב"ש ש- $L(M) \in \mathrm{DTIME}\Big(o\big(f(n)\log f(n)\big)\Big)$, כלומר קיים *DTM* M' שמכריע את L בזמן $o\big(f(n)\log f(n)\big)$. נניח שהקידוד של M' הוא מספיק ארוך. נריץ את $M(\langle M'\rangle)$. הזמן שלו הוא פחות מ- $f(n)\log f(n)$, והפלט שלו הפוך מהפלט של M', סתירה.

מסקנות מ-*THT*

1. לכל שתי פונקציות $f(n), g(n)$ שהן *time constructable* כך ש- $f(n) \cdot \log n = o\big(g(n)\big)$, מתקיים:
$$\mathrm{DTIME}\big(f(n)\big) \subsetneq \mathrm{DTIME}\big(g(n)\big)$$
2. לכל $1 \le a < b \in \mathbb{R}$, מתקיים $\mathrm{DTIME}(n^a) \subsetneq \mathrm{DTIME}\big(n^b\big)$.
3. $\mathrm{P} \subsetneq \mathrm{EXPTIME}$, כי היחס בין זמן פולינומי לזמן מעריכי הוא יחס של o.
4. לפחות אחת מתוך הטענות הבאות נכונות:
   a. $\mathrm{P} \neq \mathrm{NP}$
   b. $\mathrm{NP} \neq \mathrm{EXPTIME}$

4 נובע ישירות ממסקנה 3. כי אנחנו יודעים ש- $\mathrm{P} \neq \mathrm{EXPTIME}$, אז *NP* לא יכול להיות שווה גם ל-*P* וגם ל-*EXPTIME*.