סיכומים לקורס סיבוכיות

פרופ' עודד רגב ופרופ' מולי ספרא, סמסטר ב' 2009

חישוביות:

משימה חישובית:

חישוב פונקציה $f:\{0,1\}^* \to \{0,1\}^*$, לרוב פונקציות בוליאניות (טווח $\{0,1\}$). בצורה שקולה נגדיר שפה: $L_f = \{x \in \{0,1\}^* | f(x) = 1\}$. ידוע גם בשם בעיות החלטה/הכרעה. מודל חישובי: אלגוריתמים בשפת תכנות כלשהי / מכונת טיורינג, אין מגבלת זיכרון.

מכונת החישוב האוניברסלית:

כל תוכנית בשפת תכנות כלשהי ניתן לייצג ע"י מחרוזת של $\{0,1\}$. תהי α תוכנית, נסמן ב-M_α את האלג' המייצג תוכנית זו.

משפט (טיורינג): **קיים** אלג' U שבהינתן $x,\alpha \in \{0,1\}^*$: $U(x,\alpha) = M_\alpha(x)$ - כלומר U מסמלץ את M_α על הקלט x. אם $M_\alpha(x)$ נתקע, גם U תתקע. רעיון ההוכחה: U עובר על התוכנית α ומסמלץ כל שורה בה. טיורינג הראה שניתן להריץ תוכנית על מחשב.

משפט טיורינג (לקיום פונ' לא חשיבה):

קיימת פונ' $UC:\{0,1\}^* \to \{0,1\}$ שלא ניתנת לחישוב ע"י אף אלגוריתם (פונ' חשיבה: פונ' f שקיים עבורה אלג' M כך שלכל קלט x: $M(x)$ עוצר ומחזיר את $f(x)$).

הוכחה:

דרך א': חישובי עוצמות – גודל קבוצת הפונקציות היא א לעומת גודל קבוצת התוכניות שהיא $\aleph_0$.

דרך ב': שיטת הלכסון: נציב בטבלה את כל האלג' $M_0, M_1, M_{00}, M_{01} \dots$ אל מול כל הקלטים $0,1,00,01, \dots$ ונגדיר את UC באופן הבא:

$UC(\alpha) = \begin{cases} 0, & M_\alpha(\alpha) = 1 \\ 1, & o/w \end{cases}$ - כלומר מחזיר תוצאה הפוכה לאלכסון, או תוצאה כלשהי (0 או 1) אם התוצאה על האלכסון מתבדרת. לפי בניה זו, אף אלג' לא יחשב נכון את UC לכל קלט. מניחים בשלילה שקיים קלט β עבורו M_β מחשב את UC, ומגיעים לסתירה מבניית UC על הקלט β.

משפט: $HALT$ אינה ניתנת לחישוב.

$HALT = \{< \alpha, x > | M_\alpha \ stops \ on \ x\}$. נניח בשלילה שקיים M_{HALT}, ונראה כיצד ניתן לחשב את UC: בהינתן קלט α, M_{UC} מריץ את $M_{HALT}(\alpha,\alpha)$. אם התוצאה לא עוצרת, נעצור ונוציא 1. אם התוצאה עוצרת, נחשבה ונחזיר את ההיפך.

פרק ראשון: סיבוכיות זמן ריצה:

נסתכל לרוב על מחלקות של שפות, כלומר של פונקציות בוליאניות.

הגדרה:

עבור פונ' $T:\mathbb{N} \to \mathbb{N}$ נגדיר $DTIME(T(n))$ בתור אוסף השפות שניתן לפתור אותן בעזרת אלגוריתם שרץ בזמן $c \cdot T(n)$ עבור קבוע c כלשהו.

המחלקה *P*:

- $P = \cup_{c \geq 1} DTIME(n^c)$: הגדרה זו תקפה לכל מודל חישובי, כגון שפת C, $JAVA$, TM וכו'.
- BPP: מחלקת אלגוריתמים הרצים בזמן פול' ומשתמשים באקראיות (שאלה פתוחה: $P = BPP$).
- מחלקת BQP (שימוש בפיסיקה קוונטית).

המחלקה *NP*:

מחלקת השפות שניתן לוודא שייכות אליהן בזמן פולינומיאלי. כלומר, שפה $L \subseteq \{0,1\}^*$ היא ב-NP אם קיים פולינום $p:\mathbb{N} \to \mathbb{N}$ ואלגוריתם M הרץ בזמן פולי' כך שלכל קלט $x \in \{0,1\}^*$: $x \in L \Leftrightarrow \exists w \in \{0,1\}^{p(|x|)}. M(x,w) = 1$. כלומר, האלג' בהינתן הקלט ועד באורך פולי' לאורך הקלט יחזיר t (1) אמ"מ x בשפה.

דוגמאות לבעיות:

- קבוצה ב"ת (*IS*) – בעיה ב-*NPC*: בהינתן גרף לא מכוון G ומספר טבעי k, האם יש ב-G קבוצה ב"ת בגודל לפחות k (קבוצת קודקודים ללא קשתות). העד יכול להיות רשימה של k קודקודים. האלג' המוודא M יעבור על העד ויבדוק את תקינותו.
- *HamPath* – בעיה ב-*NPC*: בהינתן גרף לא מכוון G, האם קיים בו מסלול המילטוני – מסלול המבקר בכל קודקוד בדיוק פעם אחת.
- *HamCycle* – בעיה ב-*NPC*: כנ"ל רק בודק קיום מעגל המילטוני.
- *MaxIS* – בעיה ב-Σ_2^P: בהינתן גרף G ומספר k, האם גודל הקבוצה הבלתי תלויה המקסימלית היא בגודל k.
- *Linear Programming* – בעיה ב-*P*: בהינתן אוסף אי-שוויונים, האם יש להם פתרון מעל הממשיים.
- *Integer Programming* – בעיה ב-*NPC*: כמו קודם, רק מעל השלמים.

- *Graph Isomorphism – בעיה ב-NP*: לא הצליחו להראות שבעיה זו ב-NPC או ב-P. בהינתן שני גרפים G, H, האם הם איזומורפיים, כלומר האם הם אותו גרף תחת שינוי שמות הקודקודים. העד יהיה פרמוטציה של שמות הקודקודים, והמוודא יבדוק את תקינות הפרמוטציה ונכונותה. אם בעיה זו ב-NPC, אז $NP = coNP$.
- *Compositness – בעיה ב-P*: בהינתן מספר N האם הוא פריק, כלומר לא ראשוני. גודל הקלט הוא $log\ N$, וביחס לגודל זה מוודאים יעילות. ב-NP בגלל שעד יכול להיות k מחלק כלשהו, שגודלו הוא $log\ N$, והמוודא בודק שהוא מחלק. ב-P בגלל:
- *Primality*: בהינתן מספר N, האם הוא ראשוני. ב-$coNP$, כי ניתן לתת מחלק k. עם הזמן גילו אלג' ב-BPP ולבסוף אלג' דטרמיניסטי ב-P.
- *Factoring – בעיה ב-NP וב-coNP*: בהינתן מספרים N,k, האם ל-N יש מחלק ראשוני בתחום $\{k, k+1, \dots, N-1\}$. ב-NP: ניתן לתת כעד מחלק ראשוני בתחום, ולבדוק בזמן פולי' שהוא בתחום, מחלק וראשוני. ב-$coNP$: ניתן לקבל רשימת מספרים $a_1, \dots, a_j$ ולוודא שמכפלתם היא N, שכל אחד מהם ראשוני ושכולם קטנים מ-k.

טענה: $P \subseteq NP \subseteq EXP = \bigcup_{c\geq 1} DTIME(2^{n^c})$

הוכחה:

- $P \subseteq NP$: בונים מוודא שמתעלם מהעד – המוודא יהיה האלג' המקורי שפותר בזמן פולי'.
- $NP \subseteq EXP$: תהי $L \in NP$ וכל המשתמע מכך. נבנה אלג' M' שמפעיל את האלג' המוודא M על כל $2^{p(|x|)}$ העדים האפשריים. אם לפחות אחת מהריצות קיבלה, נקבל. אחרת נדחה.

הגדרה חלופית ל-NP:

ההגדרה החילופית משתמשת באי-דטר'. נגיד ששפה $L \subseteq \{0,1\}^*$ שייכת ל-$NTIME(T(n))$ אם קיים $c > 0$ וקיים אלג' לא דטר' M כך שלכל x: קיימת בחירה אי דטרמיניסטית שגורמת ל-M לקבל $x \in L \Leftrightarrow$. בנוסף M עוצר אחרי $c \cdot T(n)$ צעדים עבור כל בחירה אי דטר'.

משפט: $NP = \bigcup_{c\geq 1} NTIME(n^c)$

הוכחה: עבור $L \in NTIME(n^c)$, נבנה אלג' מוודא: מריץ את האלג' האי דטר' המקורי, אך במקום פיצולים אי דטר', מתפצל לפי העד הניתן לו (אלג' זה דטר'). עבור $L \in NP$, נבנה אלג' לא דטר': מריץ את האלג' המוודא תוך כדי שבונה עד באופן לא דטרמיניסטי באורך $p(|x|)$.

תרגול:

סגירות לפעולות:

נגדיר $L^* = \{y \mid \exists k \in \mathbb{N}.\, y = y_1 \dots y_k, \forall i \in \{1, \dots, k\}.\, y_i \in L\}$.

תרגיל: האם NP סגורה ל-*?

פתרון: כן. עבור $L \in NP$, נגדיר ל-x קלט ל-L^* עד באופן הבא: $0 < i_1 < i_2 \dots < i_k = n$ חלוקה ל-x ו-$w_1, \dots, w_k$ עדים לכל אחד ממחרוזות בחלוקה של x. האלג' המוודא ישתמש במוודא של L לבדוק שכל קטע בחלוקה הוא מחרוזת ב-L. העד באורך $O(n \cdot p(n))$ (p אורך עד ב-L) וזמן הריצה גם כן. נכונות ברורה.

תרגיל: האם P סגורה ל-*?

פתרון: כן. עבור מילה $x = x_1 x_2 \dots x_n$ נבנה גרף בו הצמתים $0,1,2, \dots, n$ ולכל קטע בחלוקה מאינדקס i ל-j נעביר קשת מכוונת (i, j). נבדוק אם בגרף יש מסלול מכוון מ-0 ל-n. זמן ריצה $O(n^2 \cdot p(n))$ - הרצת האלג' המכריע את L על כל הקטעים האפשריים, בדיקת מסלול מכוון ע"י BFS. נכונות ברורה.

תרגיל: האם $coNP$ סגורה ל-*?

פתרון: בהינתן קלט x באורך n יש להוכיח כי ב-G_x אין מסלול מכוון מ-0 ל-n. העד: אוסף זוגות הצמתים שאין ביניהם קשת ועדות לכך שאין ביניהם קשת ($\bar{L} \in NP$ ולכן קיימים עדים כאלה). נקח את הגרף השלם ונסיר ממנו את אותן קשתות, ונבדוק שאכן אין עליו מסלול מכוון מ-0 ל-n.

רדוקציות ושלמות ב-NP:

רדוקציית *Karp*:

פולי' משפה A לשפה B היא פונ' f חשיבה בזמן פולי' (בגודל הקלט) כך שלכל x: $x \in A \Leftrightarrow f(x) \in B$. נסמן $A \leq_p B$.

- אם $A \leq_p B$ וגם $B \leq_p C$ אז $A \leq_p C$.
- אם $A \leq_p B$ ו-$B \in P$ אז $A \in P$ או $B \in NP$ אז $A \in NP$ בהתאם.

NP-hard: A תהיה NP-קשה אם לכל $L \in NP$: $L \leq_p A$.

NP-Complete: A תהיה NPC אם היא NP-קשה וגם $A \in NP$.

משפט: $TMSAT \in NPC$

$TMSAT = \{< \alpha, x, 1^x, 1^t > \mid$ קיים $u \in \{0,1\}^n$ כך שהאלגוריתם M_α מוציא 1 על הקלט $< x, u >$ תוך זמן $t\}$. הסיבה שמשתמשים בייצוג אונארי הוא כדי שהאלג' ירוץ בזמן פולי' ל-n ולא ל-$\log n$.

<u>הוכחה בקיצור</u>: $TMSAT \in NP$: בונים מוודא המקבל u כעד ומסמלץ את M_α על $< x, u >$ למשך t צעדים. גודל הקלט וזמן הריצה פולי' ($n + t \geq$). $TMSAT \in NP - hard$: תהי $L \in NP$ עם אלג' מוודא M_α הרץ בזמן $p(|x|)$ עם עד בגודל $q(|x|)$. הרדוקציה $L \leq_p TMSAT$: תתרגם קלט x ל-L ל: $< \alpha, x, 1^{q(|x|)}, 1^{p(|x|)} >$.

משפט *Cook-Levin*: $SAT \in NPC$

<u>טענה</u>: לכל פונק בוליאנית מעל l משתנים $f: \{0,1\}^l \to \{0,1\}$ יש נוסחת CNF שקולה בגודל $\leq l \cdot 2^l$.

<u>הוכחה</u>: $SAT \in NP$: ברור. $\forall L \in NP. L \leq_p SAT$: תהא M מ"ט לא דטר' המחשבת את L ורצה בזמן $p(|x|)$, עם שתי פוני' מעברים δ_0, δ_1. נתרגם קלט x ל-L לנוסחה φ_x כך ש-$x \in L \Leftrightarrow \varphi_x \in SAT$. לכל קלט x נבנה נוסחה המתארת טבלה בגודל $p(|x|) \times p(|x|)$ המתארת את ריצת המכונה M על x – כל שורה בטבלה היא שלב. לייצוג כל תא נדרש למקום קבוע: $log\,|Q| + log\,|\Sigma| + 1$. φ_x תורכב מתנאים המוודאים שכל השורות מורכבות באופן תקין (השורה הראשונה עם q_0 מעל התא הראשון, השורה האחרונה מייצגת מצב קבלה, מעברי השורות – מצבים – נכונים ע"פ δ_0 או δ_1). סה"כ יתקבל מס' פולי' של תנאים מעל מס' קבוע של משתנים – סה"כ נוסחת φ_x בגודל פולי' והרדוקציה פולי'. יש להוכיח אמ"מ.

טענה: $IS \in NPC$

<u>הוכחה</u>: $IS \in NP$: ברור. $NP - hard \ni E3SAT \leq_p IS$ ($E3SAT = exactly\ 3\ literals\ in\ each\ clause$): לכל φ עם m הסגרים ניצור גרף עם m משולשים שקודקודי כל משולש הם הליטרלים בהסגר. בנוסף כל ליטרל יחובר לשלילתו. אם הנוסחה ספיקה אז יש בגרף קבוצה ב"ת בגודל m – לוקחים את אחד המשתנים המספקים בכל הסגר. אם יש קבוצה ב"ת בגודל m, חייב להיות בה קודקוד אחד בכל שלישיה ולכן ישנה השמה מספקת ע"י סיפוק הליטרלים כפי שמופיעים בקבוצה הב"ת.

טענות נוספות: $Clique, VC \in NPC$.

<u>תרגול</u>:

רדוקציות *Cook / Karp*:

<u>תרגיל</u>: להראות רדוקציית $Karp$: $HamPath \leq_p HamCycle$.

<u>פתרון</u>: בהינתן G הרדוקציה תחזיר G' כך ש-$V(G') = V \uplus \{u\}$ ו-$E(G) = E \uplus (\{u\} \times V) \uplus (V \times \{u\})$ (מוסיפים קודקוד ומחברים אותו לכולם).

נכונות: אם ב-G מסלול המילטוני $v_1 \to \cdots \to v_n$, ב-G' מעגל המילטוני $u \to v_1 \to \cdots \to v_n \to u$. אם ב-$G'$ מעגל המילטוני, מסירים את u ומקבלים מסלול המילטוני.

טענה: NP סגורה לרדוקציות $Karp$.

<u>תרגיל</u>: להראות רדוקציית $Cook$: $HamPath \underset{cook}{\leq} HamCycle$.

<u>פתרון</u>: בהינתן G' קלט לבעיית מעגל המילטון, נבדוק האם $E' = \emptyset$. אם כן, נחזיר שאין מעגל המילטון. אם לא, נבחר קשת כלשהי (u, v) ונריץ את $HamPath(V \uplus \{w, z\}, E \uplus \{(z, v), (u, w)\})$. אם מחזיר כן, נחזיר כן – כי יחד עם הקשת (u, v) ניתן להשלים מעגל המילטון. אחרת, נסיר את הקשת הזו ונחזור לשלב ההתחלה.

טענה: NP סגורה לרדוקציית $Cook$ פולי' אמ"מ $NP = coNP$.

<u>הוכחה</u>: אם NP סגורה לרדוקציית $Cook$ פולי', והרי לכל $L: L \in NP$ $\bar{L} \underset{cook}{\leq} L$ ע"י החזרת התשובה ההפוכה לתשובת האלג' המכריע את L, נקבל כי $\bar{L} \in NP$ ולכן $NP \subseteq coNP$ וגם $coNP \subseteq NP$ ולכן $NP = coNP$. אם $NP = coNP$, תהיה $L \in NP = coNP$ ו-$L' \underset{cook}{\leq} L$, נראה $L' \in NP$. יהי A אלג' המכריע את L'. בהינתן קלט x לשפה L', יהי k מספר הקריאות של A לאלג' המכריע את L. העד שנצפה לו הוא התשובות $a_1, \dots, a_k \in \{0,1\}$ של האלג' המכריע את L, וכן עדים לנכונות תשובות $L: w_1, \dots, w_k$. כיוון ש-$L \in NP = coNP$, ניתן לוודא האם קלט שייך או לא שייך ל-L באמצעות עד. האלג' של L' יסמלץ את A ובכל קריאה לאלג' הפותר את L יוודא בעזרת w_i אם a_i התשובה הנכונה. אם לאורך כל החישוב העדים ענו על הדרישה ו-A קיבל, אז נקבל. אחרת נדחה.

<u>חיפוש לעומת החלטה</u>:

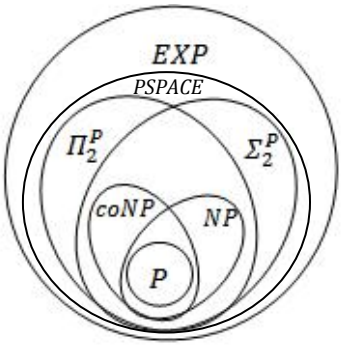
משפט: אם $P = NP$ אז לכל $L \in NP$ יש אלג' פולי' שבהינתן $x \in L$ מוציא עבורו עד.

<u>הוכחה על SAT</u>: נסמן את האלג' הפולי' הפותר את SAT ב-A. בהינתן φ נבדוק האם ספיקה. אם לא – נעצור. אם כן, נקבע את x_1 ל-t ע"י הוספת (x_1) לנוסחה ונבדוק האם ספיקה. אם כן, נמשיך עם $x_1 = t$, אחרת נמשיך עם $x_1 = f$ (נוסיף את $(\overline{x_1})$ במקום (x_1)). לבסוף מתקבלת השמה מספקת. נדרשת הוכחת נכונות.

<u>הוכחה כללית</u>: תהא $L \in NP$ עם אלג' מוודא M. נתרגם את M למ"ט דטר' M' המחשבת את L. לפי $C\text{-}L$ בהינתן x ניתן לבנות φ_x כך ש-$x \in L \Leftrightarrow \varphi_x \in SAT$, ולפי הוכחת $C\text{-}L$, בהינתן השמה מספקת ל-φ_x ניתן בזמן פולי' לשחזר את הבחירות הלא דטר' ומהן לבנות את העד שגורם ל-M לקבל.

<u>המחלקה $coNP$</u>:

$coNP = \{L \mid \bar{L} \in NP\}$. הגדרה חלופית: קיים פולינום p ומ"ט פולי' M כך שלכל $x \in \{01\}^*$: $x \notin L \Leftrightarrow \exists v \in \{0,1\}^{p(|x|)}.M(x,v) = 1$, או באופן שקול: $x \in L \Leftrightarrow \forall v \in \{0,1\}^{p(|x|)}.M(x,v) = 1$. NP ידועה גם כ-$\exists P$, ו-$coNP$ ידועה גם כ-$\forall P$.



דוגמא: $CNF - EQUIV = \{< \varphi, \psi > \mid \forall x \in \{0,1\}^*.\varphi(x) = \psi(x)\}$.

טענה: $\overline{SAT} \in coNPC$ (תחת רדוקציית $Karp$)

טענה: $CNF - EQUIV \in coNPC$

$\overline{SAT} \leq_p CNF - EQUIV$: $\varphi \rightarrow < \varphi, (x) \wedge (\bar{x}) >$ - הנוסחה השניה תמיד לא ספיקה. נכונות ברורה.

<u>המחלקות $EXP, NEXP$</u>:

$EXP = \bigcup_{c \geq 1} DTIME(2^{n^c})$, $NEXP = \bigcup_{c \geq 1} NTIME(2^{n^c})$ (ניתן להגדיר גם באמצעות עד).

ברור שמתקיים: $P \subseteq NP \subseteq EXP \subseteq NEXP$.

משפט: אם $EXP \neq NEXP$ אז $P \neq NP$

<u>הוכחה</u>: נניח $P = NP$. ברור כי $EXP \subseteq NEXP$, נראה $NEXP \subseteq EXP$. תהי $L \in NEXP$, נגדיר בשיטת ה-***Padding***: $L_{pad} = \{< x, 1^{2^{|x|^c}} > \mid x \in L\}$. $L_{pad} \in NP$: בהינתן קלט נבדוק שהוא מהצורה הנדרשת, ואם כן נפעיל עליו את האלג' הלא-דטר' הרץ בזמן $O(2^{|x|^c})$. זמן הריצה הוא פולי' בגודל הקלט ב-L_{pad}. אם כן, לפי הנחה, $L_{pad} \in P$, ומכאן קל לבנות אלג' אקספ' דטר' המכריע את L.

<u>המחלקות Σ_2^P, Π_2^P</u>:

<u>Σ_2^P</u>: מחלקת השפות שיש עבורן מ"ט פולי' M ופולינום p כך שלכל $x \in \{0,1\}^*$: $x \in L \Leftrightarrow \exists u \in \{0,1\}^{p(|x|)}.\forall v \in \{0,1\}^{p(|x|)}.M(x,u,v) = 1$.

דוגמא: $MIN - DNF$ - בהינתן נוסחת DNF φ ומספר k, האם קיימת נוסחת DNF ψ בגודל לכל היותר k כך ש-$\psi \equiv \varphi$: $\exists \psi. \forall x. \varphi(x) = \psi(x)$ (בעיה Σ_2^P-שלמה).

<u>Π_2^P</u>: מוגדרת בצורה דומה עם החלפת ה-$\forall$ עם ה-$\exists$. מתקיים: $L \in \Sigma_2^P \Leftrightarrow \bar{L} \in \Pi_2^P$.

דוגמא: $MAX - IS = \{< G, k > \mid G'smax.ind.set\ is\ exactly\ of\ size\ k\}$. ניתן להביע את התנאי בשתי הצורות הנ"ל, שייך ל- $\Sigma_2^P \cap \Pi_2^P$, הקרויה גם DP. בעיה זו היא DP-שלמה.

ניתן להמשיך ולהוסיף תנאי $\exists, \forall$ ולהגדיר Σ_3^P, Π_3^P וכן הלאה, והיררכיה זו קרויה ההיררכיה הפולינומיאלית – PH.

משפט: $P = NP \Leftrightarrow \Sigma_2^P = P$

<u>הוכחה</u>: כיוון אחד ברור. אם נניח ש-$P = NP$, ותהי $L \in \Sigma_2^P$ אז $x \in L \Leftrightarrow \exists u \in \{0,1\}^{p(|x|)}.\forall v \in \{0,1\}^{p(|x|)}.M(x,u,v) = 1$. נגדיר $L' = \{< x, u > | u \in \{0,1\}^{p(|x|)}, \forall v \in \{0,1\}^{p(|x|)}.M(x,u,v) = 1\}$. מהגדרה $L' \in coNP = P$, ונניח M' הוא האלג' המוודא של L'. מכאן: $x \in L \Leftrightarrow \exists u \in \{0,1\}^{p(|x|)}.M'(x,u) = 1$. באופן דומה זו שפה ב-$NP$ ולפי הנחה ב-P.

משפט הירכית הזמן:

לכל פונקציות f,g כך ש-(g גדולה משמעותית יותר מ-f) $g(n) = \omega(f(n) \cdot log\, f(n))$ מתקיים: $DTIME(f(n)) \subsetneq DTIME(g(n))$.

<u>הוכחה</u>: נגדיר בעית החלטה: בהינתן אלג' α האם הפעלת U (המ"ט האוניברסלית) לסמלץ את M_α על α למשך $g(|\alpha|)$ צעדים עוצרת? מהגדרה, הבעיה ב-$DTIME(g(n))$, נראה שאינה ב-$DTIME(f(n))$. חסר.

<u>תרגול</u>:

<u>3-צביעה</u>:

- בעיית הכרעה: בהינתן גרף G, האם ניתן לצבוע את צמתיו בשלושה צבעים $\{1,2,3\}$ כך שכל שני צמתים סמוכים צבועים שונה?
- בעיית חיפוש: בהינתן גרף G, מצא צביעה לקודקודיו בשלושה צבעים $\{1,2,3\}$ כך שכל שני צמתים סמוכים צבועים שונה.

תרגיל: למצוא רדוקצייה עצמית ($Cook$) מבעיית החיפוש לבעיית ההכרעה (כלומר: בהינתן אלג' מכריע ל-$3\text{-}col$, למצוא אלג' המוצא צביעה).

פתרון: בהינתן G תחילה נבדוק האם $G \in 3-col$, אם לא אז נדחה. אם כן, נוסיף לו שלושה צמתים חדשים $\{x,y,z\}$ שנחברם במשולש ונסמן גרף זה G'. לכל $v \in V$ נפעיל את אלג' ההכרעה על שלושה גרפים עם צמתי G' וקשתות G', כאשר לראשון נוסיף הקשתות $\{v,x\},\{v,y\}$, לשני $\{v,x\},\{v,z\}$ ולשלישי $\{v,y\},\{v,z\}$. מתוך הגרפים האלה נבחר גרף כלשהו G_i $(1 \le i \le 3)$ שהוא 3-צביע ונסמנו G'. הצבע של הצומת v יהיה i כאשר G_i הוא הגרף שנבחר מהאיטרציה על v. נכונות: G' 3-צביע ולכן ל-v צביעה בצבע מבין 1,2,3, ולכן הוספת 2 קשתות לצמתים מבין x,y,z שאינם צבועים בצבע זהה, תשאיר את הגרף 3-צביע לפי אותה צביעה.

תרגיל: אם קיימת שפה אונארית NP-שלמה אז $P=NP$.

הוכחה: תהא $L \subseteq \{1\}^* \ni NPC$, ולכן $SAT \le_p L$. תהא f הרדוקציה שלכל φ מחזירה $f(\varphi)=1^i$ כאשר $i \le p(n)$ כאשר p פולינום ו-n מספר משתני הנוסחה φ / אורך φ. נראה אלג' ל-SAT הרץ בזמן פולי' ומכך נסיק $P=NP$. נשתמש במערך A שכל ערכיו מאותחלים ל-$unknown$. נגדיר את $SAT(\varphi(x_1,\dots,x_n),A)$:

- אם $n=0$ החזר את φ (t או f).
- אם $A\left[|f(\varphi)|_{=i}\right] \neq unknown$, החזר את $A[|f(\varphi)|]$.
- אם $SAT(\varphi(t,x_2,\dots,x_n),A)$ או $SAT(\varphi(f,x_2,\dots,x_n),A)$ נשים t ב-$A[|f(\varphi)|]$ ונחזיר t, אחרת נשים f ונחזיר f.

הרעיון: עבור קלט בגודל n יתכנו 2^n קלטים אפשריים, אך f ממפה אותם לקבוצה קטנה יחסית – בגודל $p(n)$ (כי לכל i הוא היחיד באורך $x \in L$ הוא 1^i ואין עוד קלטים אחרים אפשריים באורך i). לכן שמירת הערכים ב-A תחסוך בדיקות בהמשך = תחסוך התפצלויות בעץ האלג' הנאיבי.

נכונות: נובעת מכך ש-$\varphi(x_1,\dots,x_n)$ ספיקה אמ"מ $\varphi(t/f,x_2,\dots,x_n)$ ספיקה. זמן ריצה: בכל שלב נבחר צומת שמתאים לקריאה רקורסיבית בעץ הרקורסיה (לא עלה) ונסיר אותו ואת המסלול המוביל אליו מהשורש: $O(n)$ צמתים. נעשה זאת עד שהעץ יתרוקן. כל קריאה רקורסיבית כזו הנמצאת בתחתית המסלול מתאימה לערך שונה מהמערך A, ולכן לכל היותר יוסרו $p(n)$ מסלולים. סה"כ: $O(n \cdot p(n))$. בכל קריאה כזו מפעילים את f ולכן סה"כ $O(p^2(n)\cdot n)$.

<u>אורקלים ומגבלת הלכסון</u>:

אורקל:

תהא $O \subseteq \{0,1\}^*$ שפה כלשהי. אלג' עם גישה לאורקל O הוא אלג' שיש לו פקודה המאפשרת לו לפתור את O בצעד אחד של קריאה לאורקל. אלג' זה יסומן M^O. בצורה דומה נגדיר אלג' לא דטר' עם גישה לאורקל ומכאן את המחלקות P^O, NP^O.

טענה: $\overline{SAT} \in P^{SAT}$. הוכחה: הפעלת האורקל על קלט φ והחזרת התשובה ההפוכה.

טענה: עבור $O \in P$ מתקיים $P^O = P$. הוכחה: $P \subseteq P^O$ - ברור. $P^O \subseteq P$ - ניתן לסמלץ את האורקל ע"י מכונה ב-P ולהישאר בזמן פולי'.

טענה: $EXPCOM = \{< \alpha, x, 1^n > | M_\alpha(x)=1, runs\ in\ 2^n\ steps\}$; $NP^{EXPCOM} = P^{EXPCOM} = EXP$

<u>הוכחה</u>:

1) <u>$EXP \subseteq P^{EXPCOM}$</u>: תהי $L \in EXP$ עם מכונה M_α הפותרת את L בזמן 2^{n^c} עבור c כלשהו. נבנה אלג' המשתמש ב-$EXPCOM$ הרץ בזמן פולי' באורך הקלט: על קלט x מחזיר את תשובת האורקל ל-$< \alpha, x, 1^{n^c} >$.

2) <u>$P^{EXPCOM} \subseteq NP^{EXPCOM}$</u>: ברור.

3) <u>$NP^{EXPCOM} \subseteq EXP$</u>: בזמן אקספ' ניתן לעבור על כל האפשרויות הלא-דטר' גם עם האורקל.

משפט *Baker-Gill-Solovay*: קיים אורקל A כך ש-$P^A \neq NP^A$

<u>הוכחה</u>: לכל שפה A נגדיר $U_A = \{1^n | \exists x \in A. |x|=n\}$. לכל A מתקיים $U_A \in NP^A$ כי העד יכול להגיד לנו איפה נמצא את המחרוזת $x \in A$. נותר לבנות A כך ש-$U_A \in P^A$.

<u>תרגול</u>:

<u>טענה</u>: $EXP^{EXP} \neq EXP$.

<u>הוכחה</u>: נראה ש-$EXP^{EXP} \subseteq DTIME\left(2^{2^n}\right)$ ואז לפי משפט היררכית הזמן: $EXP \subsetneq EXP^{EXP}$. תהא $L \in DTIME(2^{2^n})$ אז קיימת מ"ט M הרצה בזמן $c \cdot 2^{2^n}$ ומכריעה את L. נראה כי $L \in EXP^{EXPCOM} \subseteq EXP^{EXP}$: בהינתן קלט x ניצור את הקלט לאורקל $< M, x, 1^{2^{n^{+c'}}} >$ ונחזיר את תשובת האורקל. נשים לב כי $\forall x.\ x \in L \Leftrightarrow < M, x, 1^{2^{n^{+c'}}} > \in EXPCOM$. זמן הריצה של M חסום ע"י 2^{2^n} ולכן $EXP \subsetneq EXP^{EXP}$.

<u>פרק שני: סיבוכיות מקום</u>:

תהי $t:\mathbb{N}\to\mathbb{N}$ פונקצית סיבוכיות מקום (מונוטונית לא יורדת).

$SPACE(t(n))$: מחלקת השפות שניתנות להכרעה במקום $O(t(n))$ ע"י מ"ט דטרמניסטית.

$NSPACE(t(n))$: מחלקת השפות שניתנות להכרעה במקום $O(t(n))$ ע"י מ"ט לא דטרמיניסטית.

המחלקות *NL ,L*:

$L=SPACE(log(n)), NL=NSPACE(log(n))$. נגדיר מ"ט בעלת שלושה סרטים באופן הבא:

1. סרט קלט: קריאה בלבד.
2. סרט עבודה: קריאה וכתיבה. גודל סרט זה קובע את סיבוכיות המקום.
3. סרט פלט: לכתיבה בלבד.

קונפיגורציות:

מספר הקונפיגורציות עבור קלט בגודל N וסרט עבודה בגודל S: $\underbrace{|Q|}_{\text{המכונה מצב}}\times\underbrace{S}_{\substack{\text{מיקום הראש}\\ \text{בסרט העבודה}}}\times\underbrace{|\Gamma|^S}_{\text{תוכן סרט העבודה}}\times\underbrace{N}_{\substack{\text{מיקום הראש}\\ \text{בסרט הקלט}}}\times\underbrace{|\Sigma|^N}_{\text{תוכן סרט הקלט}}$

דוגמאות לשפות ב-L: ... $a^nb^n, a^nb^na^n, a^nb^{2n}a^n, palindrome, a^nb^{2n}a^{4n}b^{8n}$.

שאלה פתוחה: האם קיימת שפה ב-$NL\backslash L$; האם $NL\subsetneq P$.

רדוקצית *log-space Karp*:

$A\leq_L B$ אם קיימת פונ' $f:\Sigma^*\to\Sigma^*$ הרצה במקום לוגריתמי כך שלכל w: $w\in A\Leftrightarrow f(w)\in B$.

משפט:

המחלקות $EXP, PSPACE, NP, P, NL, L$ סגורות תחת רדוקצית מקום לוגריתמי.

רדוקציה לוגריתמית במקום מ-A אל B: מסמלצים ריצה של f (סרט ה-out לא חסום במקום, לכן ניתן לסמלץ צעד צעד, ולשמור על שימוש במקום לוגריתמי).

בעיית *CONN*:

בהינתן גרף G וקודקודים s,t, האם קיים מסלול מ-s ל-t. $CONN\in NL$:

להלן אלג' במקום לוגריתמי לא דטר': יהי $u=s$ קודקוד התחלתי נבצע $|V|$ איטרציות: בכל שלב u בוחר שכן כלשהו וממשיך אליו. אם $u=t$, נקבל, אחרת נמשיך. אם לא הגענו ל-t, נדחה. את u ואת i (מספר הצעדים שביצענו עד כה, עד שנגיע ל-$|V|$) מחזיקים ב-$log\ |V|$ מקום.

***NL-TM*:**

1. סרט קלט: קריאה בלבד.
2. סרט עבודה: קריאה וכתיבה.
3. סרט עדות: לקריאה בלבד, רק משמאל לימין, ללא אפשרות לחזור אחורה.

$CONN\in NL-Complete$:

בהינתן מכונה M וקלט x ניצור את $G,\ s,\ t$: נגדיר קונפיגורציה מקבלת: נמחק את סרט העבודה, ונגדיר קונ' זו כמקבלת. $G_{M,x}$ יהיה גרף הקונפיגורציות, שכל קודקוד בו יתאים לקונפיגורציה של M על הקלט x. מס' הקונ' הוא פולי' באורך x. קשת (u,v) תהיה ב-G אם יש מעבר $u\to v$ לפי δ (פונ' מעברים לא דטר') ב-M, כלומר מעבר קונפיגורציות חוקי. s יהיה הקונ' ההתחלתית ו-t תהיה הקונ' היחידה (כפי שהוגדרה לעיל) המקבלת.

$G_{M,x}$ ניתן ליצירה במקום לוגריתמי: נחזיק את הגרף במט' שכנויות התופסת מקום לינארי, אך נעבוד על קשת אחת בכל זמן נתון לפני הדפסתה לסרט הפלט – ולכן נזדקק לסרט עבודה לוגריתמי בלבד.

נותר להוכיח כי $\forall M,x: M\ accepts\ x\Leftrightarrow\exists s\to t\ in\ G_{M,x}$.

משפט *Savitch*:

$\forall S(n)\geq log(n): NSPACE\big(S(n)\big)\subseteq SPACE(S^2(n))$: כל מכונה המשתמשת במקום $S(n)$ לא דטר', ניתן לסמלץ ע"י מכונה דטר' עם $S^2(n)$ מקום. להלן אלג' דטר' ל-$CONN$ המשתמש ב-$log^2(n)$ מקום:

נראה אלג' הבודק האם בגרף מכוון G קיים מסלול מ-u אל v באורך לכל היותר d, ואז נבדוק על $s,t,d=|V|$: אם $(u,v)\in E$, נקבל. אחרת, אם $d=1$, נדחה. אחרת, נבדוק האם לכל $w\in V$ מתקיים האלג' על $u,w,\left\lceil\frac{d}{2}\right\rceil$ ועל $w,v,\left\lfloor\frac{d}{2}\right\rfloor$. אם כן, נקבל. אחרת נדחה. עומק הרקורסיה באלג' הוא $log\ |V|$. בכל רמה של הרקורסיה צריך לשמור w – דורש מקום לוגריתמי בלבד. סה"כ: $log^2|V|$ מקום.

<u>תרגול</u>:

$2SAT \in coNL - Complete$:

<u>$\overline{2SAT} \in NL$</u>:

בהינתן $\varphi \in 2CNF$ בעלת n משתנים, נבנה גרף עם $2n$ משתנים (לכל משתנה ושלילתו). נייצג את האילוצים בגרף ע"י קשתות מתאימות באופן הבא: כדי שהסגר כלשהו יסופק צריך להתקיים שאם שלילת הליטרל הראשון מקבל t, אז הליטרל השני צריך לקבל t. כמו כן אם שלילת הליטרל השני מקבלת t, הליטרל הראשון צריך לקבל t (כדי שיהיה לפחות ליטרל אחד מתוך ה-2 בהסגר שמקבל t). לכל הסגר $(x_1 \vee x_2)$ נסמן אם כן בגרף את הקשתות $(\overline{x_1}, x_2)$ ו-$(\overline{x_2}, x_1)$.

<u>טענה</u>: ל-φ **אין** השמה מספקת $\Leftrightarrow$ קיים משתנה x_i שיש מסלול ממנו ל-$\overline{x_i}$ (ומסלול הפוך גם כן). אם הטענה נכונה, קל להראות ש-$\overline{2SAT} \in NL$: ננחש באופן אי דטר' את x_i, וננחש באופן לא דטר', כמו באלג' המכריע את $CONN$, מסלול מ-x_i אל $\overline{x_i}$ ולהיפך. אלג' זה עובד עם מקום $log\ (n)$ לא דטר'.

<u>הוכחת הטענה</u>:

אם קיים i כך שיש מסלול מ-x_i אל $\overline{x_i}$ ולהיפך, אז לפי הגדרת הגרף כל צומת שניתן להגיע אליו מ-x_i, אם $x_i = t$ אז גם אותו צומת מקבל t. לפיכך נקבל כי גם $\overline{x_i} = t$, וכמו כן אם $\overline{x_i} = t$ אז $x_i = t$, וזו סתירה. לפיכך ל-φ אין השמה מספקת.

נניח כעת כי לכל i לא קיים מסלול מ-x_i אל $\overline{x_i}$. כאמור, אם $x_i = t$, כל x_j שיש מסלול מ-x_i אליו מקבל t. נראה שלא יתכן שנקבל סתירה. נניח כי במתן ערכי אמת ל-x_i, x_k קיבלנו סתירה לערכו של x_j. מכאן שקיימים מסלולים מ-x_i ל-x_j ומ-x_k ל-$\overline{x_j}$. לפיכך, יש גם מסלול מ-x_j ל-$\overline{x_k}$, ולכן מסלול מ-x_i אל $\overline{x_k}$. לכן, כאשר קבענו את ערכו של x_i להיות t, קבענו בכך את ערכו של x_k להיות f, ומכאן שלא היתה סתירה בקביעת x_j.

<u>$CONN \leq_L \overline{2SAT}$</u>:

בהינתן G,s,t נגדיר את φ באופן הבא: $\varphi = (s \vee s) \wedge (\bar{t} \vee \bar{t}) \wedge_{(u,v) \in E} (\bar{u} \vee v)$. אם קיים מסלול בין s ל-t, אז כדי ש-φ תסופק, s חייב להיות t ו-t חייב להיות f וכל צומת שיש מסלול מ-s אליו חייב להיות t. לכן, לא יתכן שיש השמה מספקת ל-φ. אם אין ב-G מסלול מ-s ל-t, ניתן ערך t ל-s ולכל המשתנים שיש מסלול מ-s אליהם. נשים לב שאין קשת מצומת u שיש מסלול מ-s אליו לצומת v שאין מסלול מ-s אליו (אחרת היה מסלול $s \to u \to v$). לכן, כל ההסגרים מהצורה $(\bar{u} \vee v)$, אם $\bar{u} = f$ אז $u = t$, ומכיוון שיש מסלול מ-s אל v אז גם $v = t$. לכן כל ההסגרים יסתפקו.

<u>הרדוקציה במקום לוגריתמי</u>: -

<u>סימולציית חישוב לא דטר' בחישוב דטר'</u>:

נבנה מכונת טיורינג דטר':

- סרט קלט.
- סרט סימולציה.
- סרט $Guide$ – אורכו בדיוק $t(n)$ (זמן הריצה של המכונה הלא דטר'). יחזיק מספר שיהווה את העדות. נסמלץ לפי העדות, ובכל פעם שלא מקבלים מעלים את העדות ב-1, כך עד מיצוי כל האפשרויות ($t(n)$ 1-ים). אם בסוף לא קיבלנו, נדחה.

<u>מקום</u>: מקום סרט הסימולציה וסרט ה-$Guide$ הוא סה"כ $O(t(n))$. סרט הסימולציה – הסימולטור רץ בזמן $t(n)$, ולכן זהו מספר התוים שלכל היותר יכתוב, והעדות – כאמור בגודל זה גם כן.

$SAT \leq_p Clique$:

רדוקציה מ-$3SAT$ באופן הבא: עבור $\varphi \in 3SAT$ בעלת m הסגרים נבנה $< G, m >$ כך שב-G:

- V: קודקוד לכל מופע של כל ליטרל, כלומר סה"כ $3m$ קודקודים.
- E: אין קשתות בין מופע של ליטרל למופע ההפוך לו, ובין כל שלישית קודקודים מאותו הסגר. כל שאר הקשתות קיימות.

<u>שלמות</u>: אם φ ספיקה, אז הקליקה מכילה לכל הסגר קודקוד אחד לפחות – זה שמספק את ההסגר (מקבל t).

<u>תקפות</u>: בהינתן קליקה ב-G בגודל k, לכל משתנה מהנוסחה או שמופיע במופע שלילי או בחיובי, אך לא שניהם יחד – כי לפי הבניה של G אין ביניהם קשת. נציב באותם משתנים בקליקה t וכך כל ההסגרים יסתפקו, כי לכל הסגר יש "נציג" בקליקה. ההצבה במשתנים שלא בקליקה לא משנה.

<u>מקום</u>: הרדוקציה היא $logspace$, מיידי.

<u>Scale-up למשפט Savitch</u>:

יהיו $s_1(n), s_2(n) \geq log\ (n)$ פונקציות זיכרון, ותהי $e(n) \geq n$ פונקצית הרחבה. אם $NSPACE(s_1(n)) \subseteq SPACE(s_2(n))$ אז $NSPACE(s_1 \circ e(n)) \subseteq SPACE(s_2 \circ e(n))$.

הוכחה:

תהי $L \in NSPACE(s_1 \circ e(n))$, נבצע $padding$: $L^e = \{x\#^{e(|x|)-|x|} | x \in L\}$ - הופכים את גודל הקלט ל-$e(|x|)$ במקום $|x|$. האלג' יבדוק שהקלט אכן מהצורה הזו (בדיקת $|x|$, חישוב $e(|x|)$ וחישוב $e(|x|) - |x|$ ולבסוף בדיקת הקלט עצמה) ולאחר מכן נפעיל את האלג' הפותר ב-$NSPACE(s_1 \circ e(n))$ את הבעיה. האלג' רץ עבור קלט בגודל $e(|x|)$ ולכן אלג' זה להכרעת L^e גורר ש-$L^e \in NSPACE(s_1(n))$.

לפי ההנחה, קיים אלג' דטר' המכריע את L^e במקום $SPACE(s_2(n))$, נסמן את המ"ט שעושה זו ב-M'. נסמלץ את מכונת M' ב-$SPACE(s_2 \circ e(n))$, באופן הבא: בהינתן קלט x בלבד, המכונה תסמלץ את פעולת M' כך שתספור כאילו $e(|X|)$ תוי #, אך למעשה פשוט תריץ מונה עד $e(|x|)$. כיוון שיש $s_2 \circ e(n)$ מקום, ו-$s_2 \geq log$, הרי שיש מספיק מקום להחזיק מונה עד $e(|x|)$. ומכאן מכונה הרצה במקום $SPACE(s_2 \circ e(n))$ המכריעה את L.

טענה: $NL = coNL$:

נראה זאת ע"י כך שנראה ש-$NON - CONN \in NL$, כלומר הבעיה: האם אין מסלול בין s ל-t, שהיא $coNL - Complete$ ב-NL.

נסתכל על מודל מכונת NL בו יש סרט עדות שניתן לקריאה משמאל לימין ללא חזרה. נבנה את העד בצורה אינדוקטיבית:

משפט *Immerman-Szelepcsenyi*: $NON - CONN \in NL$

נגדיר את הקבוצה $reachable(G) = \{v | \exists s \to v\}$ - קבוצת כל הקודקודים הנגישים מ-s. נשאל האם $t \in reachable(G)$. נסתכל על $G_{-t} = (V \backslash \{t\}, E(V \times \{t\}))$, שהוא גרף ללא t, ונשאל: $\#reachable(G) \stackrel{?}{=} \#reachable(G_{-t})$. אם כן, אזי $< G, s, t > \in NON - CONN$.

מספיק להראות ש-$\#reachable(G) = r$ עבור r כלשהו.

נגדיר: $reachable_l(G) = \{v | \exists s \to v \ of \ length \leq l\}$. אינדוקציה על l כשלבסוף $l = |V|$ יכסה את כל $reachable(G)$.

בסיס: s - $r_0 = 1$ הקודקוד הנגיש לעצמו ב-0 צעדים.

מעבר: נניח יש לנו עדות ל-$r_l = \#reachable_l(G)$, נראה הרחבת העדות ל-r_{l+1}. העדות מהצורה $\#r_l\#w$, ונרחיבה: לכל $1 \leq i \leq |V|$ נשרשר לעדות עד כה את $\$w_i$ כאשר הביט מסמל האם הקודקוד i שייך ל-$reachable_{l+1}(G)$ (1 אם כן, 0 אם לא), ו-w_i העדות לכך. בדיקת השייכות ל-$reachable_{l+1}(G)$:

- שייכות: w_i תהיה מסלול באורך $\geq l + 1$ שהוא $s \to i$ אותו יוודא האלג'. אם המסלול הניתן לא טוב, נדחה. אם כן נוסיף 1 למונה r_{l+1}.
- אי שייכות: w_i תהיה עדות לכך שכל הקשתות שנכנסות לקודקוד i אינן מתחילות מקודקוד ששייך ל-$reachable_l(G)$. העדות תהיה אם כן: לכל $1 \leq j \leq |V|$ ישורשר z_j *< ביט >, כאשר הביט מייצג האם הקודקוד j לא נגיש מ-s ב-l צעדים, ואם כן – שאין קשת (j, i). לשלב זה נחזיק מונה שמתחיל מ-0:
  - אם j נגיש מ-s ב-l צעדים: z_j יהיה מסלול באורך $\geq l$ שהוא $s \to j$, ונבדוק שאין קשת (j, i). אם יש קשת כזו, נדחה. אחרת נעלה את המונה ב-1.
  - אם j לא נגיש מ-s ב-l צעדים: z_j תהיה עדות ריקה.

  כל מה שצריך כדי לוודא שזה נכון הוא לבדוק בסוף שמספר הנגישים שספרנו לפי עדויות ה-z_j השונות הכולל (המונה שהחזקנו) שווה בדיוק ל-r_l שחושב קודם לפי הנחת האינדוקציה. אם לא – נדחה. אם כן – נמשיך.

לבסוף, לאחר מעבר על כל $1 \leq i \leq |V|$ נוודא שעבור G, G_{-t} התקבל אותו $r_{|V|}$. אם כן – נקבל, ואחרת נדחה.

תרגול:

לאיזו מחלקת סיבוכיות שייכת הבעיה:

תרגיל: בהינתן גרף G, האם יש בו לכל היותר 2009 רכיבי קשירות חזקים: ***NL***.

פתרון: מנחשים (האלג' מנחש באופן אי דטר') קבוצה $S \subseteq V$ של 2009 קודקודים. לכל $v \in V$: מנחשים $s \in S$ ומנחשים מסלול בין s ל-v וחזרה ל-s. אם נמצאו מסלולים כנ"ל לכל $v \in V$, נקבל. אחרת, נדחה. הנכונות מכך שאם G בשפה, קיים ניחוש אי דטר' שמוצא 2009 נציגים לכל אחד מ-2009 רכיבי הקשירות, כך שלכל קודקוד נוכל למצוא את הנציג שלו, ונוודא זאת ע"י מציאת מסלול אליו וחזרה. אם בגרף יותר מ-2009 רכיבי קשירות, אז קיים v כך שלכל $s \in S$ שנבחר מתוך 2009 הקודקודים ב-S, אף אחד לא נציג של v, ולכן לא נמצא מסלולים $s \to v, v \to s$.

תרגיל: האם ב-G לפחות 2009 רכיבי קשירות חזקים: ***NL***

פתרון: בעיה זו היא הבעיה המשלימה של התרגיל הקודם, שזו בעיה ב-NL. לפי $Immerman$, $NL = coNL$, ולכן גם בעיה זו ב-NL.

תרגיל: ב-G בדיוק 2009 רכיבי קשירות: ***NL***

פתרון: בעיה זו חיתוך של השתיים הקודמות, ו-NL סגורה לחיתוך ($L_1, L_2 \in NL \Rightarrow L_1 \cap L_2 \in NL$).

תרגיל: בעיית $CONN$ לגרפים לא מכוונים: $\boldsymbol{L}$ – הוכח ע"י $Reingold,\ 2005$.

תרגיל: בהינתן G מכוון עם n צמתים, האם קיים מעגל פשוט באורך $\frac{n}{2}$: $\boldsymbol{NPC}$

פתרון: $HamCyc \leq_p \left(\frac{n}{2}\right) SimpleCyc$: יוצרים G' שהוא הכפלת הגרף G. ב-G' מעגל פשוט באורך $\frac{n}{2}$ אמ"מ ב-G מעגל המילטוני (פשוט באורך n).

תרגיל: בהינתן G גרף מכוון, האם קיימים $s,t \in V$ כך שאורך המסלול הקצר ביותר שמחבר אותם הוא באורך $\leq \frac{n}{2}$: $\boldsymbol{NL}$

פתרון: הבעיה המשלימה: האם כל שני צמתים במרחק $< \frac{n}{2}$ - בעיה ב-NL: עוברים על כל זוגות הצמתים ומנחשים מסלול באורך כזה.

תרגיל: בהינתן גרף G, האם קיימת בו IS שהיא I ו-$Clique$ שהיא C כך ש-$|I| + |C| \geq 2009$: $\boldsymbol{L}$

פתרון: יש פתרון גם בזמן קבוע: בכל גרף עם לפחות $\binom{4016}{2008}$ צמתים יש קליק או IS בגודל 2009. הפתרון כאן: עוברים על כל i ובודקים האם יש קליק בגודל i ו-IS בגודל $2009 - i$, כל זה במקום לוגריתמי.

<u>בעיות שלמות ב-$PSPACE$</u>:

בעיות שניתן לפתור במקום פולי'. ידוע ש-$PH \subseteq PSPACE$, אך <u>לא יודעים להוכיח</u> $P \neq PSPCAE$ (אם היה שוויון PH היתה קורסת).

טענה: השפה $SPACETM = \{< \alpha, w, 1^n > |$ מקום $n - M_\alpha(w) = 1\} \in PSPACE - Complete$. המחלקה $PSPACE$ סגורה לרדוקציה במקום לוגריתמי, וזו הרדוקציה בה נשתמש (רדוקציה במקום פולי' חזקה מדי).

נוסחה מכומתת לחלוטין – *Truly Quantified Boolean Formula (TQBF)*:

נוסחה מהצורה: $\forall/\exists x_1 \ldots \forall/\exists x_n . \varphi(x_1, \ldots, x_n)$ כאשר φ נוסחה ללא כמתים.השפה $TQBF$ – שפת הנוסחאות הללו שהן נוסחאות אמת. SAT היא תת שפה של $TQBF$ – הכמתים תמיד ישיים והנוסחאות הן מצורת CNF.

טענה: $TQBF \in PSPACE$:

<u>הוכחה</u>: להלן פרוצדורה לפתרון נוסחה φ בגודל m ב-$TQBF$: נקרא בצורה רקורסיבית לנוסחה פעם אחת עם הצבת $x_1 := t$ ופעם עם $x_1 := f$.

- אם הכמת הוא $\exists$, נקבל אמ"מ <u>לפחות אחת</u> מהקריאות מחזירה t.
- אם הכמת הוא $\forall$, נקבל אמ"מ <u>שתי</u> הקריאות מחזירות t.

<u>נכונות</u>: קל לראות נכונות, והזיכרון בו משתמשים הוא n - $O(m \cdot n)$ עומק הרקורסיה ו-$O(m)$ זיכרון להצבת ערכים ב-x_1.

מסקנה: $PH \subseteq PSPACE$

השוני בין $PSPACE$ ל-PH הוא שב-PH מספר החלפות הכמתים ידוע מראש, בעוד שב-$PSPACE$ הוא תלוי הנוסחה המתקבלת כקלט.

<u>הערה</u>: לא מאמינים ש-$PH \subseteq TQBF$ כי מספר החלפות הכמתים ב-$TQBF$ יכול להיות n ולא קבוע כמו ב-PH.

משפט: $TQBF \in PSPACE - Complete$:

נראה ש-$TQBF \in NP - hard$: בהינתן $L \in SPACE(s(n))$ עבור פולינום s כלשהו המחושבת ע"י M, לכל x נבנה ψ נוחסת $TQBF$ כך שהיא נוסחת אמת אמ"מ $x \in L$. נסתכל על $G_{M,x}$ גרף הקונפיגורציות של ריצת $M(x)$, ויהי $m = s(n)$ מספר הביטים הנדרשים לייצוג קודקוד בגרף (מספר הקודקודים $\geq 2^m$). נבנה ψ הבודקת מסלול באורך $\geq 2^m$ מקודקוד c_{start} לקודקוד c_{accept}:

נסמן את $\varphi(c_1, c_2)$ כנוסחה הבודקת האם יש מסלול בגרף $G_{M,x}$ מ-c_1 אל c_2 (מעבר חוקי של קונ'), והיא בגודל פולי'. נגדיר רקורסיבית $\psi_i(c,d)$ כנוסחה האומרת: קיים מסלול חוקי בין c ל-d באורך 2^i, ונקח לבסוף את $\psi_m(c_{start}, c_{accept})$: $\psi(c_1, c_2) = \exists d . \psi_{i-1}(c_1, d) \wedge \psi_{i-1}(d, c_2)$. נוסחה זו מתפתחת אקספ', ולכן כדי לקצרה נהפוך כל זוג הופעות של ψ_{i-1} בהופעה בודדת: $\exists d . \underbrace{(\forall (d_1, d_2) \in \{(c_1, d), (d, c_2)\})}_{\equiv \forall (d_1,d_2)((d_1 = c_1 \wedge d_2 = d) \vee \ldots) \rightarrow} . \psi_{i-1}(d_1, d_2)$, כאשר החלק המסומן הוא מגודל קבוע.

<u>פרק שלישי: אלגוריתמי קירוב, בעיות אופטימיזציה</u>:

<u>אלגוריתמי קירוב</u>:

אלגוריתם קירוב ל-VC בפקטור 2:

<u>האלג'</u>: בונים קבוצת זיווג מקסימלי $M \subseteq E$, נוסיף לקבוצה זו קשתות שלא נוגעות בקשתות שכבר בקבוצה עד שלא ניתן. המקסימלי יכול להשתנות בהתאם לבחירת הקשתות, אין זה משנה כאן. נגדיר את ה-VC: $C = \{v \in V | \exists e \in M \ s.t. \ v \in e\}$ - קבוצת הקודקודים של קשתות M.

<u>זמן ריצה</u>: ברור שפולי', כיוון שמציאת זיווג מקסימלי נעשית בזמן פולי'.

<u>נכונות</u>: אכן מתקבלת קבוצת כיסוי חוקית כיוון שכל קשת ב-E חייבת לגעת ב-M כיוון ש-M מקסימלית.

יחס הקירוב ≥ 2: מתקיים $|OPT(G)| \leq |C| \leq 2|OPT(G)|$: גודל הפלט הוא $|C| = 2 \cdot |M| = |ALG(G)|$. מכיוון ש-$M$ זיווג, גודל האופטימום הוא לפחות M כי האופטימום מכיל לפחות קודקוד אחת מכל קשת ב-M. כלומר: $|M| \leq |OPT(G)|$, ולכן $|C| \leq 2|OPT(G)|$.

הערה: ידוע שזה NP-קשה לקרב ביחס 1.36 וכמו כן שזה קשה (לא NP-קשה) לקרב עד כדי $2 - \varepsilon$.

בעיות תכנות לינארי: בעיות LP ניתנות לפתרון פולינומיאלי בעזרת אלגוריתם האליפסואיד.

אלגוריתם קירוב ל-VC בעזרת LP:

בהינתן קלט ל-VC נכתוב את תוכנית האופטימיזציה הבאה:

כאשר $x_i = \begin{cases} 1, x_1 \in VC \\ 0, \quad o/w \end{cases}$ קודקוד. בגלל שהאילוץ האחרון הוא לינארי, בעיה זו היא בעיית *Integer Programming (IP)* והיא NPC.

Minimize $\sum_{i=1}^{n} x_i$ *subject to:*

- $\forall \{i,j\} \in E. x_i + x_j \geq 1$ – לכל קשת לפחות קודקוד אחד בכיסוי
- $\forall i. x_i \in \{0,1\}$

נקח בעיה מקורבת המחליפה את האילוץ האחרון באילוץ לינארי: $\forall i. x_i \in [0,1]$, כלומר הקלנו בתנאים. נסמן את האופטימום המקורי ב-OPT ואת האופטימום של הבעיה המקורבת (בעיית ה-LP) ב-$OPT*$. ברור כי $OPT \leq OPT*$. בנוסף $OPT*$ ניתנת לחישוב בזמן פולי' עם האליפסואיד.

האלג': נמצא פתרון $x_1, \dots, x_n \in [0,1]$ כך ש-$\sum x_i = OPT*$, ונגדיר את הכיסוי: $C = \{i | x_i \geq 0.5\}$.

זמן ריצה: פולי' בעזרת האליפסואיד.

נכונות: הפלט הוא כיסוי חוקי של צמתים כי לכל קשת $\{i,j\}$: $x_i + x_j \geq 1$ ולכן לפחות אחד מהם ≤ 0.5, ולכן קשת זו מכוסה.

ניתוח קירוב: $|C| \leq 2 \sum_{i=1}^{n} x_i = 2OPT* \leq 2OPT$ - כי כל 1 ב-C תורם לפחות 0.5 ב-$OPT*$ - יחס קירוב גם כן 2.

תרגול:

בעית מקסימיזציה: אלג' A מהווה c-קירוב אם לכל x ערך הפתרון ש-A מספק מקיים: $A(x) \geq \frac{OPT(x)}{c}$.

בעית מינימיזציה: אלג' A מהווה c-קירוב אם לכל x ערך הפתרון ש-A מספק מקיים: $A(x) \leq c \cdot OPT(x)$.

בעיית TSP:

בהינתן גרף מלא $G = K_n$ עם פונ' משקלות $w: E \to \mathbb{R}$, יש למצוא מעגל פשוט העובר בכל קודקוד הגרף בעל משקל מינימלי.

טענה: בעיה זו לא ניתנת לקירוב לכל קבוע c שהוא. נראה 2-קירוב לבעית TSP כאשר w מקיימת אשמ"ש: $w(x,y) + w(y,z) \geq w(x,z)$:

2-קירוב ל-TSP עם אשמ"ש:

הוכחה: בהינתן G נמצא עפ"מ. לאחר מכן נכפיל כל קשת בגרף כולו – G, ונמצא מעגל אויילר בעץ המוכפל. נלך על המעגל, ובכל פעם שאמורים לבקר בצומת בו היינו, נדלג דרך קשת מוכפלת לצומת בו טרם היינו. בצומת האחרון נקפוץ לצומת ההתחלה ע"י קשת מוכפלת.

נכונות: משקל מסלול מינימלי $\leq$ משקל עפ"מ, ולכן $MST \leq 2 \cdot TSP_{\text{אשמש}} \leq 2 \cdot$. כל קפיצה במהלך המסלול לא מעלה את משקל המסלול הכולל בגלל האשמ"ש על w, ולכן האלג' הולך במסלול השוקל לכל היותר $2 \cdot MST$.

$\frac{3}{2}$-קירוב ל-TSPעם אשמ"ש:

בהינתן גרף H שידוך מושלם לגרף הוא $E' \subseteq E$ כך שכל צומת מופיע בדיוק בקשת אחת ב-E'.

טענה: בגרף מולא עם מספר זוגי של צמתים ניתן למצוא שידוך מושלם עם משקל מינימלי בזמן פולי'.

טענה: בכל גרף, אם O קבוצת צמתים מדרגה אי זוגית אז $|O|$ זוגי.

האלג': נמצא עפ"מ ב-G, נסמנו T. עבור O קבוצת הצמתים מדרגה אי זוגית נמצא שידוך מושלם במשקל מינימלי בתת הגרף $G[O]$ ונסמנו M. ב-$T \cup M$ יש דרגה זוגית לכל צומת. נבנה עליו מעגל אויילרי ונפעל כמו האלג' הקודם (הכפלת קשתות וכו').

ניתוח קירוב: $MST \leq TSP$. משקל סיור TSP אופטימלי ב-$O \leq 2 \times$משקל שידוך מינימלי ב-O (כי סיור ניתן להפריד לשני שידוכים – עבור מעגל קשתות $e_1 - e_2 - \cdots e_n - e_1$ נקח את כל הקשתות הזוגיות לשידוך אחד והאי זוגיות לשני. לפי אשמ"ש משקל TSP ב-$G \leq$ משקל TSP ב-$G[O] \leq 2 \times$ משקל M. מכאן: $T \cup M \leq \frac{3}{2} \cdot TSP$.

בעיות אופטימיזציה:

דוגמאות: $Clique$ – מקסימיזציה (מציאת קליקה גדולה ביותר), SAT – מקסימיזציה (רוצים לספק כמה שיותר פסוקיות), VC – מינימיזציה (מציאת כיסוי קודקודים קטן ביותר). לכל בעיה קלט ופרמטר אותו בודקים כמדד. למשל: $Clique$ - $|clique|$, SAT - $\#clauses$, VC - $|cover|$.
האלג' צריך להחזיר פתרון חוקי שמתוכם את הכי טוב.

הגדרה: קירוב-α: אלגוריתם לבעית מקסי'/מיני' עם פתרון אופטימלי O, המחזיר פתרון $\leq \frac{O}{\alpha}$ (מקסי') או $\geq \alpha \cdot O$ (מיני').

בעית *Set-Cover*:

בהינתן קבוצת איברים U ו-$F \subseteq P(U)$ כך ש-$\bigcup_{A\in F} A = U$ רוצים למצוא $C \subseteq F$ מינימלית המכסה את כל U: $\bigcup_{A\in C} A = U$. זו בעית מינימיזציה NP-קשה: $VC \leq_p Set - Cover$: בהינתן גרף G, נבנה אינסטנציה ל-$SetCover$ באופן הבא: ב-U יהיה אלמנט לכל קשת ב-E, וב-F תהיה קבוצה לכל קודקוד. כמו שכל קודקוד מתאים לקשתות, כך כל קבוצה מ-F מתאימה לאלמנטים ב-U.

אלג' קירוב לבעית *Set-Cover* – אלג' חמדן:

מתחילים מ-C קבוצה ריקה ו-U' המאותחלת ל-U. כל עוד $U' \neq \emptyset$: ניקח $S \in F$ כך ש-$|S\cap U'|$ מקסימלי, כלומר S המכסה את מקסימום האלמנטים שנותרו שטרם כוסו. נוסיף את S ל-C. חמדנות האלג' נובעת מכך שמבצע את מה שהכי טוב בכל שלב בנפרד.

יהי k האופטימום, כלומר כל תת קבוצה של U (ו-U עצמה) ניתנת לכיסוי ע"י k תתי-קבוצות. <u>מסקנה</u>: בכל צעד נלקחת קבוצה המכסה לפחות $\frac{1}{k}$ מהאלמנטים שנותרו.

<u>הקירוב הוא $log_2 n$</u>: לאחר k צעדים מכוסים לפחות $\frac{1}{2}$ מהקודקודים, ולכן לאחר $log_2 n \cdot k$ צעדים תכוסה כל U. מדוע לאחר k צעדים מכוסים לפחות $\frac{1}{2}\cdot n$ קודקודים: נניח בשלילה שלא, אז לאחר k צעדים נשארים $> \frac{n}{2}$ קודקודים לא מכוסים. כיוון שבכל שלב נלקחים $\frac{1}{k}$ לפחות מהנותרים, הקבוצה הבאה שתלקח תהיה $> \frac{n}{2k}$. מצד שני פו' מספר הקודקודים שנלקחים בכל שלב לא עולה מונוטונית (בכל שלב נלקחים פחות ופחות פריטים), ולכן כל שלב כיסה $\leq \frac{n}{2k}$ קודקודים, וסה"כ לאחר k צעדים ראשונים: $k \cdot \frac{n}{2k} = \frac{n}{2}$.

<u>קירוב $ln\, n$</u>: יהיו $S_1, \dots, S_t$ הקבוצות שנבחרו, U_i קבוצת הקודקודים שנותרו לאחר i צעדים. מתקיים: $|U_{i+1}| = |U_i - S_{i+1}| \leq |U_i| \cdot \left(1 - \frac{1}{k}\right)$, והחסם העליון הוא כיוון שהורדנו לפחות $\frac{1}{k}$ מהקודקודים בהורדת S_{i+1}. מכאן: $|U_{ik}| \leq \underbrace{|U_0|}_{=U} \cdot \left(1 - \frac{1}{k}\right)^{ik} \leq |U| \cdot e^{-i} \Rightarrow t \leq k \cdot ln\, n + 1$.

<u>חסם עליון טוב ביותר: $H(max\{|S||\, S \in F\})$</u>: $H(n) = \sum_{i=1}^{n} \frac{1}{i} \leq ln\, n + 1$, כלומר הפקטור הוא ln של גודל הקבוצה המקסימלית ב-F: כל קבוצה S שמכניסים ל-C תורמת 1 לגודל C, והיחס בין מספר זה ל-k הוא פקטור הקירוב. נחלק את ה"1" שכל קבוצה בגודל L תורמת על פני כל האלמנטים בה, כאשר כל אחד תורם $\frac{1}{L}$. במקרה הגרוע, כל אלמנט משלם בתורו $\frac{1}{\text{גודל נוכחי}}$ כלומר סה"כ הקבוצה משלמת $H(L) = \frac{1}{1} + \cdots + \frac{1}{L-1} + \frac{1}{L}$.

סה"כ ישולמו (קבוצה מקסימלית) $k \cdot H$.

<u>בעיות *GAP*</u>:

תהי A בעית אופטימיזציה. נגדיר סף כך שכל הפתרונות מעל הסף טובים, ומתחתיו לא טובים: בעית $Gap - A[C, hC]$. בעית GAP קלה מהבעיה המקורית ולכן $GAP - A \in NP - hard \Rightarrow A \in NP - hard$.

טענה: לכל C, אלג' h-קירוב ל-A מבטיח פתרון לבעית $Gap - A[C, hC]$:

- <u>מינימיזציה</u>: נגדיר שלכל קלט x, אם $A(x) < hC$ נקבל, אחרת נדחה (אם $A(x) < hC$, מובטח ש-$OPT(x) < C$, ולכן נקבל).
  בעית $GAP[C, hC]$ מינימיזציה: מקבלים לפלטים $< C$, דוחים לפלטים $< hC$.
- <u>מקסימיזציה</u>: נגדיר שלכל קלט x, אם $A(x) > \frac{C}{h}$ נקבל, אחרת נדחה (אם $A(x) > \frac{C}{h}$, מובטח ש-$OPT(x) > C$, ולכן נקבל).
  בעית $GAP[C, hC]$ מקסימיזציה: מקבלים לפלטים $< hC$, דוחים לפלטים $< C$.

<u>מסקנה</u>: אם בעית ה-GAP היא NP-קשה, כך גם בעית הקירוב (אלא אם $P = NP$).

<u>$GAP - 3SAT[\frac{7}{8} + \varepsilon, 1]$</u>:

קלט טוב: 100% מההסגרים מסתפקים; קלט לא טוב: פחות מ-$\frac{7}{8} + \varepsilon$ מהסגרים מסתפקים. נראה שלכל נוסחת $3CNF$ יש השמה המספקת $\frac{7}{8}$ מההסגרים: בכל הסגר C_i בו בדיוק 3 משתנים שונים נסמן משתנה אינדיקציה y_i שמקבל 0 אם C_i לא מסתפקת מהצבה רנדומית כלשהי ו-1 אם כן. לכל הסגר כזה המקרה היחיד שלא יסופק הוא עבור השמה f לכל משתניו, בסיכוי $\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{8}$. לכן $E[y_i] = \frac{7}{8}$. אם בנוסחה m הסגרים, מלינאריות של התוחלת: $E[\sum y_i] = \sum E[y_i] = \frac{7}{8}\cdot m$, כלומר $\frac{7}{8}$ מהסגרי הנוסחה יסופקו. ההוכחה לכך שיש הצבה כזו: כיוון שהממוצא הוא $\frac{7}{8}$ מההסגרים מסתפקים, חייבת להיות הצבה שלפחות כמו הממוצע.

<u>משפט ה-*PCP*</u>:

לכל $\varepsilon > 0$: $Gap - 3SAT\left[\frac{7}{8} + \varepsilon, 1\right]$ היא NP-קשה. כך ניתן לבצע רדוקציה מכל $L \in NP$:

$x \notin L \Rightarrow max\ sat\ \frac{7}{8} + \varepsilon\ clauses$; $x \in L \Rightarrow f(x) \in 3SAT$.

<u>תרגול</u>:

<u>דוגמא</u>: נניח $Gap-IS[\frac{1}{4},\frac{1}{3}]$ היא NP-קשה, אזי קשה לקרב את IS בפקטור $\frac{4}{3}=\frac{1}{3}/\frac{1}{4}$.

<u>תרגיל</u>: **בעית $max-2SAT$**:

ידוע ש-$2SAT$ פולי', ולכן לכל $0\le a<1$ הבעיה $Gap-2SAT[a,1]$ פולי'. רדוקציה מ-$Gap-E3SAT[\frac{7}{8}+\varepsilon,1]$: כל הסגר מהצורה $(x\vee y\vee z)$ יועבר ל-10 הסגרים עם w משתנה חדש מהצורה: $(x)\wedge(y)\wedge(z)\wedge(w)\wedge(\bar{x}\vee\bar{y})\wedge(\bar{x}\vee\bar{z})\wedge(\bar{y}\vee\bar{z})\wedge(\bar{w}\vee x)\wedge(\bar{w}\vee y)\wedge(\bar{w}\vee z)$.

בהינתן השמה ל-x,y,z:

- אם היא לא מספקת אף אחד מ-x,y,z, אז לכל בחירה של w לא יותר מ-6 מההסגרים יסופקו בעשיריה החדשה.
- אם היא מספקת ליטרל אחד בדיוק, לא יותר מ-7 הסגרים יסופקו בהשמה החדשה.
- אם 2 ליטרלים מסופקים נקבל 7 הסגרים מסופקים לכל בחירה של w.
- אם כל הליטרלים מסופקים נקבל לכל היותר 7 הסגרים מסופקים.

אם במקור היו m הסגרי $3CNF$, אז אם יכולנו לספק את כולם, נוכל לספק בעיה החדשה $7m$ מתוך $10m$ הסגרים. אם לא יכולנו לספק יותר מ-$\left(\frac{7}{8}+\varepsilon\right)m$ מההסגרים המקוריים, לא נוכל לספק יותר מ-$\left(\frac{55}{8}+\varepsilon\right)m$-$m\cdot 6+\left(\frac{1}{8}-\varepsilon\right)\cdot 7m+\left(\frac{7}{8}+\varepsilon\right)$ הסגרים מתוך $10m$ הסגרים. לפי משפט ה-PCP: $Gap-2SAT\left[\frac{55}{80}+\varepsilon,\frac{7}{10}\right]\in NP-hard$ לכל $\varepsilon>0$, ומסקנה: קשה לקרב את $max\text{-}2SAT$ בפקטור $\frac{56}{55}-\varepsilon=\left(\frac{55}{80}+\varepsilon\right)/\frac{7}{10}$.

<u>וריפיקציה פרובביליסטית</u>:

טענה: $Gap-IS[1-\beta,1-\alpha]\equiv Gap-VC[\alpha,\beta]$.

נשתמש במכונת טיורינג עם סרט קלט, סרט עבודה וסרט עדות. כל הוכחה ניתנת לרדוקציה לנוסחת $3CNF$, וסרט העדות יהיה הצבה מספקת. נבחר באופן רנדומי $O(1)$ ביטים (הצבות) מסרט העדות, למשל 100, כאשר ידוע לפי משפט ה-PCP שהסיכוי לספק הסגר מסוים כאשר ההצבה שניתנה (העדות) <u>אינה</u> מספקת, הוא $\frac{7}{8}+\varepsilon$. ואז הסיכוי לטעות בבדיקה ולזהות הצבה לא מספקת כמספקת הוא: $\left(\frac{7}{8}+\varepsilon\right)^{100}$ - סיכוי קטן מאוד.

משפט: $Gap-Clique\left[\left(\frac{7}{8}+\varepsilon\right)m,m\right]\in NP-hard$:

הקלט: גרף בו m קבוצות ב"ת בגודל 3. הפרמטר הוא הקליקה המקסימלית בגרף. במקרה זה קלטים טובים הם כאלה עם בדיוק m קודקודים בקליקה המקס'. בעיה זו אינה NP כיוון שאינה בעית הכרעה, אך היא NP-קשה. נראה זאת ע"י <u>רדוקציה משמרת Gap</u> מ-$3SAT$:

<u>רדוקציה משמרת Gap</u>: מעבירה קלטים טובים לטובים, רעים לרעים ו-$don't\text{-}cares$ לא משנה לאן.

כל הסגר הופך לשלושה קודקודים לא מחוברים ביניהם ואין קשתות בין ליטרלים הפוכים, כל שאר הקשתות קיימות. m מספר ההסגרים.

<u>נכונות</u>: אם יש השמה מספקת לכל היותר $\frac{7}{8}+\varepsilon$ מההסגרים, הקליקה המקסימלית היא בגודל לכל היותר $\frac{7}{8}+\varepsilon$. אם יש השמה מספקת לכל ההסגרים, הקליקה היא בגודל m.

<u>Graph Constraints Problems – CSG</u>:

קלט: $U=(V,E,\Sigma,\phi)$ כאשר V,E גרף, Σ קבוצת צבעים ו-$\phi:E\to P(\Sigma^2)$ פונקציית אילוצי צביעה על קשתות הגרף. שתי אפשרויות:

- CSG_V: גרסת הקודקודים, נותנים השמה (צביעה) $A:V\to\Sigma\cup\{\perp\}$ כך שלכל $A(u),A(v)\in\Sigma$ מתקיים $<A(u),A(v)>\in\phi(u,v)$. המטרה היא למקסם את כמות הצמתים שמקבלים צבע מ-Σ (ולא $\perp$).
- CSG_E: גרסת קשתות: $A:V\to\Sigma$, המטרה למקסם כמות הקשתות $(u,v)\in E$ המקיימות את האילוצים: $<A(u),A(v)>\in\phi(u,v)$.

<u>אופטימיזציה</u>: בגרסת הקודקודים נרצה למקסם את $Pr[A(u)\in\Sigma]$ - ההסתברות שקודקוד יקבל צבע מ-Σ ולא $\perp$.

רדוקציה מ-$max-3SAT$:

לכל הסגר יהיה קודקוד, $\Sigma=\{1,2,3\}$ (שלושה צבעים בדיוק). עבור שני הסגרים $c_1=(a\vee b\vee c),c_2=(x\vee y\vee z)$ (כאשר אלו <u>ליטרלים</u>), אם למשל $a=\neg y$ אז $(1,2)\notin\phi(c_1,c_2)$, המייצג שאת c_1 הליטרל ה-1 (a) מספק, ואת c_2 הליטרל ה-2 (y) מספק, שכן זו תהיה סתירה.

Max-Cut: רדוקציה לבעית CSG עם שני צבעים, צבע לכל קבוצה בחתך, ולכל $(u,v)\in E$ יתקיים $\phi(u,v)=\{(1,2),(2,1)\}$.

טענה:

$Gap-3CSG_V\left[\frac{7}{8}+\varepsilon,1\right]$ היא NP-קשה: עולה ישירת מהרדוקציה משמרת ה-GAP מ-$Gap-3SAT$ לעיל.

טענה:

$Gap-kCSG_V[\delta,1]\le_L Gap-IS[\frac{\delta}{k},\frac{1}{k}]$.

הוכחה: לכל $v \in V$ יהיו k קודקודים $v_1, \dots, v_k$ (קודקוד לכל צבע), כאשר כל הקודקודים האלו מחוברים בקליקה (לכל קודקוד $v \in V$ קליקה משלו) – $V' = V \times \Sigma$. שאר הקשתות בגרף יוגדרו:

- לכל $i \neq j$ אז $(u_i, u_j) \in E'$ - קליקה עבור כל הקודקודים החדשים לכל קודקוד מהגרף המקורי.
- אם $(i,j) \notin \phi(u,v)$ אז $(u_i, v_j) \in E'$ - אם הצביעה (i,j) אינה חוקית ל-(u,v), נעביר קשת (u_i, v_j) בגרף G'.

עבור צביעה הצובעת חלק מקודקודי G, ב-G' יש קשת היכן שהאילוץ לא מתקיים – היכן שאין קשת ב-G. בהינתן IS ב-G' נמצא צביעה חוקית ל-G: נצבע קודקוד u שעבורו ב-IS ב-G' נמצא הקודקוד u_i בצבע i. ה-GAP חלקי k כיוון שגודל הקבוצה נשאר זהה, בעוד גודל הגרף גדל פי k.

משפט:

$Gap - kCSG_V[\delta, 1] \leq_L Gap - k^l CSG_V[\delta^l, 1]$.

מסקנה מהמשפט: בעית IS לא ניתנת לקירוב ע"י פקטור קבוע כלשהו.

הוכחה: הרדוקציה: נגדיל את V ואת Σ פי l. כל l-ית קודקודים ב-U היא קודקוד אחד ב-U'. האילוצים:

- כל קודקוד u המופיע בכל l-יה יהיה באותו צבע (u לא יופיע פעם אחת בצבע i ופעם אחרת בצבע j).
- האילוצים המקוריים יהיו תקפים גם בין l-יות שונות.

ב-U' יש k^l צבעים כי צביעת כל קודקוד למעשה צובעת l-יה שלמה.

נכונות: שלמות – אם ב-U מתקיימים כל האילוצים על כל הקודקודים, ניתן ב-U' לכל l-יה את הצביעה המקורית, וכל האילוצים יתקיימו. תקפות – אם יש צביעה הצובעת ε' מהקודקודים, נייצר צביעה ל-U באופן הבא: כל האילוצים מתקיימים על כל ε' מהקודקודים, ולכן כל קודקוד שמופיע ב-ε' נצבע לפי צביעת U' אותם. כל שאר הקודקודים לא יצבעו. יהי ε אחוז הצביעה שהתקבל כתוצאה מכך ב-U, נבדוק את היחס ε'/ε: מתקיים $\varepsilon' \leq \varepsilon^l$ כי ε' חסום ע"י הגודל היחסי של כל ה-l-יות הכוללות את ε אחוז הקודקודים. ומכאן, אם $\varepsilon = \delta$ אז $\varepsilon' = \delta^l$. דרך נוספת לראות זאת: אם יש לנו ε אחוז צבוע ב-U, הסיכוי של כל l-יה רנדומית להצבע הוא $\varepsilon^l = \varepsilon \cdot \dots \cdot \varepsilon$ – ε-סיכוי לכל קודקוד להיות ב-l-יה.

מסקנה: כיוון שמספר הצבעים קבוע למרות שגדל, קיבלנו GAP קרוב ל-1 כרצוננו.

תרגול:

בעיות Gap – המשך:

טענה: אם $Gap - IS[\alpha, \beta]$ היא NP-קשה, אז:

1. $Gap - Clique[\alpha, \beta]$ היא NP-קשה.
2. $Gap - VC[1-\beta, 1-\alpha]$ היא NP-קשה.

הוכחה:

1. נשתמש ברדוקציה משמרת-GAP הבאה: $G = (V, E) \to G' = (V, V^2 \backslash E)$. נכונות מיידית: IS בגודל k ב-G אמ"מ $Clique$ בגודל k ב-G'.
2. טענה: S היא IS אמ"מ $V \backslash S$ היא VC. אם S היא IS ומניחים כי $V \backslash S$ אינה VC, אז קיימת קשת (u,v) שאותה לא מכסה, ולכן $u, v \in S$, אז זו סתירה לכך ש-S היא IS. אם $V \backslash S$ היא VC, אז לכל קשת לפחות אחד מקודקודיה ב-$V \backslash S$. לכן, לא תתכן קשת (u,v) כך ש-$u, v \in S$ ולכן S היא IS. לכן הרדוקציה: בהינתן $< G, k >$ קלט לבעית ה-IS, נחזיר $< G, n-k >$ קלט לבעית ה-VC.

דוגמאות למסקנות: $Gap - VC\left[\frac{8}{10}, \frac{9}{10}\right]$, $Gap - IS\left[\frac{1}{10}, \frac{2}{10}\right]$ הן NP-קשות, קשה לקרב את VC בפקטור $\frac{9}{8} - \varepsilon = \frac{9}{10} / \frac{8}{10}$ לכל $\varepsilon > 0$.

תרגיל: $Gap - IS\left[\frac{1}{2}, \frac{4}{5}\right] \in P$.

הוכחה: הבעיה שקולה לבעיה $Gap - VC\left[\frac{1}{5}, \frac{1}{2}\right]$ שהיא ב-P, היות וקיים אלג' 2-קירוב לבעית VC.

השיטה ההסתברותית:

Max-Cut: טענה: בכל גרף קיים חתך שגודלו לפחות $\frac{|E|}{2}$.

הוכחה: כל צומת יושם ב-V_1 או V_2 בהסתברות חצי, כל בחירה ב"ת באחרות. נסמן לכל $e \in E$ משתנה X_e שיקבל 1 אם $e \in CUT(V_1, V_2)$ ו-0 אחרת. מתקיים $E(X_i) = \frac{1}{2}$, ולפי לינאריות התוחלת: $E[|CUT|] = E[\sum_{e \in E} X_e] = \sum_{e \in E} E[X_e] = |E| \cdot \frac{1}{2}$. כיוון שזהו ממוצע, קיימת הגרלה בה גודל החתך הוא לפחות $\frac{|E|}{2}$.

בעיות CSG:

בעית צביעת גרפים: בהינתן G ופרמטר k, האם קיימת ל-G צביעה חוקית ב-k צבעים: תיאור באמצעות CSG_V: $\Sigma = \{1, \dots, k\}$. צביעה חוקית (הגדרת ϕ): כל צביעה בה זוג הצמתים מקבלים ערכים שונים. המטרה: למקסם צביעה חוקית ולא $\perp$ לצמתים.

MaxCut: ייצוג בעזרת CSG_E: $\Sigma = \{1,2\}$, לכל $e \in E$ צביעה חוקית היא כזו בה שני הצמתים מקבלים צבע שונה אחד מהשני.

IS: ייצוג בעזרת CSG_V: $\Sigma = \{1\}$, לכל קשת מתקיים $\phi(e) = \emptyset$, כלומר אף זוג מעל $\{1\}$ אינו צביעה חוקית. 1 ינתן לצמתים ב-IS ו-$\perp$ לשאר.

Clique: הפוך מ-IS.

בעיות $qCSG_\Delta$:

תהי $\Delta: E \to P([q])$. האילוצים יקבעו בבעיות אלו ע"פ ההפרש בין ערך ה-q של הקודקודים. האילוצים: $\phi(u,v) = \{(i,j) | i - j \bmod q \in \Delta\}$.

טענה: $Gap - kCSG_V[\delta, 1] \leq_L Gap - (nk)^5 CSG_\Delta[\delta, 1]$, כאשר $q = (nk)^5$.

הרדוקציה: יהי מספר הקודקודים בגרף ה-CSG_V m. נבנה גרף חדש בו $m \cdot q$ קודקודים. לכל שני קודקודים, אם נוסיף לשניהם d, ההפרש ישמר ולכן ימשיכו לקיים את האילוצים בגרף ה-CSG_Δ, ויצרו IS נוספת. אם אין קשת (i,j) אז גם לא תהיה קשת $(i+d, j+d)$ ולהיפך. לכן לכל קבוצה ב"ת שנוסיף לצבעי הקודקודים בה d, נקבל קבוצה ב"ת חדשה.

מסקנה: מכאן שבעיה זו NP-קשה, ולכן בעית צביעה (חלוקת גרף ל-k קבוצות IS) היא NP-קשה.

תוצאה:

$Gap - \chi\left[q, \frac{q}{\delta}\right] \in NP - hard$: כאשר χ הוא מספר הצביעה של הגרף (המס' המינימלי הדרוש לצבוע את הגרף צביעה חוקית).

הוכחה: $Gap - qCSG_V[\delta, 1] \leq_L Gap - \chi\left[q, \frac{q}{\delta}\right]$: לוקחים את ההצבה A לקודקודי V, ומגדירים $A^d(v) = A(v) + d, \forall d$. שמירת האילוצים מתקיימת כיוון שההפרשים נשארים זהים. אם בגרף המקורי ניתן לצבוע את כל הקודקודים, ברור דגם בגרף החדש. אם היה ניתן לצבוע לכל היותר δ מהקודקודים, בגרף החדש ניתן יהיה לצבוע לכל היותר $\delta \cdot m$ מהקודקודים. סה"כ הצביעה: $\chi(G) = \frac{\#\{v \in V\}}{max(IS\ in\ G)} = \frac{q \cdot m}{\delta \cdot m} = \frac{q}{\delta}$.

ייחודיות בשלשות:

עבור $q = |U|^5$ ניתן לבנות באופן יעיל $T: U \to \{0,1, \dots, q-1\}$ כך ש:

$\forall u_1, \dots, \forall u_6: T(u_1) + T(u_2) + T(u_3) = T(u_4) + T(u_5) + T(u_6) \Rightarrow \{u_1, u_2, u_3\} = \{u_4, u_5, u_6\}$. כלומר, פונ' הנותנת לכל קודקוד מספר מתוך q מספרים, כך שלכל שלישית קודקודים סכום ייחודי לה, וניתן לשחזר ממנה את הקודקודים המשתתפים בה.

נניח שנתנו מספרים ל-l קודקודים. בשלב ה-$l+1$: כדי לתת לקודקוד ה-$l+1$ צבע מספר חוקי, צריך לבדוק לכל זוג מ-l הקודמים שהמספר שניתן לו בצירוף עם הזוג, אינו מפר סכום כל שלישיה מתוך ה-l. לפיכך נזדקק ל-l^5 בדיקות בשלב זה, וזהו מספר המספרים המקסימלי שנפסל.

הערה: ניתן לזהות בודדים וזוגות ע"י השלמתם לשלישיה וזיהוי המשתתפים בה. ולכן T ייחודית בבודדים, זוגות ושלישיות.

חזרה להוכחת הטענה: $Gap - kCSG_V[\delta, 1] \leq_L Gap - (nk)^5 CSG_\Delta[\delta, 1]$

נבנה $U' = (V, E = V \times V, \{0,1, \dots, q-1\}, \Delta)$ כך שלכל קודקוד u_i מתאימים צבע מ-0 עד $q-1$, כך שכל שלישיה יחודית בסכומה. הגודל q נקבע לפי הוכחת חסם למציאת T, והוא $q = (|V| \times k)^5 = (nk)^5$.

נגדיר את האילוצים: $\Delta(u,v) = \{T(u,i) - T(v,j) \bmod q \mid (i,j) \in \phi(u,v)\}$ - אם (i,j) מקיים את אילוצי (u,v) בגרף המקורי, נאפשר את ההפרש $T(u,i) - T(v,j) \bmod q$ ל-(u_i, v_j) בגרף שבונים.

שלמות: אם אפשר לצבוע את כל הגרף המקורי ע"י הצביעה $A: V \to \{1, \dots, k\}$, אז בגרף החדש תתאים הצביעה $A'(v) = T(v, A(v))$.

תקפות: נוכיח שלכל A' צביעה של U', יש הזזה d כך ש-$A'(u) = T(u, A(u)) + d \bmod q$ כאשר A היא צביעה מתאימה ל-U. הדרך היחידה לקבל צביעה ל-U' היא לקחת צביעה של הגרף המקורי, להפעיל T על הקודקוד והצבע שלו ולהזיז את כולם ב-d כלשהו – וכאן בדיוק הפעולה ההפוכה. כעת ידוע שמתקיים: $A'(u) - A'(v) \in \Delta(u,v)$, ורוצים להגיד שקיים (i,j) זוג צבעים יחיד כך ש-$T(u,i) - T(v,j) = A'(u) - A'(v) \bmod q$. מהסבר ארוך בסופו של דבר מגיעים לכך שמייחודיות T נובעת תקפות (עמודים 10-9 ב-*PCP writeup*).

תרגול:

המשך בעיות CSG:

רדוקציה 1: $Gap - kCSG_V[\delta, 1] \leq Gap - IS\left[\frac{\delta}{k}, \frac{1}{k}\right]$. **מסקנה**: קשה לקרב את בעית *Max-IS* לכל פקטור קבוע.

רדוקציה 2: אמפליפיקציה: $Gap - kCSG_V[\delta, 1] \leq_p Gap - k^l CSG_V[\delta^l, 1]$ לכל $l \geq 1$.

תרגיל: נניח $Gap-4CSG_V\left[\frac{1}{2},1\right]\in NP-hard$. איזו תוצאת קושי קירוב מתקבלת עבור VC? האם משתפר בעזרת אמפליפיקציה?

פתרון: $Gap-4CSG_V\left[\frac{1}{2},1\right]\leq_p Gap-IS\left[\frac{1}{8},\frac{1}{4}\right]\leq_p Gap-VC\left[1-\frac{1}{4},1-\frac{1}{8}\right]=Gap-VC\left[\frac{3}{4},\frac{7}{8}\right]$. מכאן שתוצאת הקירוב שקשה לקרב את VC יהיה פקטור של $\frac{7}{8}\cdot\frac{4}{3}-\varepsilon=\frac{7}{6}-\varepsilon$ לכל $\varepsilon>0$.

אם היינו משתמשים באמפליפיקציה: $Gap-4^lCSG_V\left[\frac{1}{2^l},1\right]\leq_p Gap-IS\left[\frac{1}{8^l},\frac{1}{4^l}\right]\leq_p Gap-VC\left[1-\frac{1}{4^l},1-\frac{1}{8^l}\right]$ – ותוצאה זו לא משפרת, כיוון שה-GAP הצטמצם ולכן מתקבל שקשה לקרב פקטור שקרוב יותר ל-1 (OPT) מאשר הפקטור שהתקבל קודם – ולכן לא נוסף מידע חדש (אם קשה לקרב ב-$\frac{7}{6}-\varepsilon$, ברור שקשה לקרב בפקטור קרוב ממנו ל-1, כפי שהאמפליפיקציה נתנה).

<u>תרגול (אותו נושא)</u>:

טענה:

יהי G גרף פשוט, $\alpha(G)$ גודל IS מקסימלית, $\chi(G)$ מספר הצביעה, n צמתים. מתקיים: $\chi(G)\cdot\alpha(G)\geq n$.

<u>CSG_Δ</u>: מקרה פרטי של CSG_V בו לקשת $(u,v)\in E$ קיימת $S_e\subseteq\mathbb{Z}_q(=\{0,1,..,q-1\})$ כך ש-(x,y) צביעה חוקית ל-(u,v) אמ"מ $x-y\in S_e\ mod\ q$.

מהטענה מההרצאה: $Gap-kCSG_V[\delta,1]\leq_p Gap-(nk)^5CSG_\Delta[\delta,1]$ נובע ש-$Gap-k'CSG_\Delta[\delta,1]$ היא NP-קשה. נתבונן ברדוקציה $Gap-kCSG_V[\delta,1]\leq Gap-IS\left[\frac{\delta}{k},\frac{1}{k}\right]$ כאשר גרף האילוצים הוא מהצורה הנ"ל.

<u>שלמות</u>: אם ניתן היה לצבוע את כל צמתי הגרף כך שכל הקשתות יהיו טובות, אז קיימת לנו בגרף הקלסטרים (החדש) קבוצה ב"ת S_0 עם נציג מכל קלסטר. אם מוסיפים i לצביעה זו מקבלים עדיין צביעה חוקית שמגדירה קבוצה ב"ת חדשה עם נציג מכל קלסטר, ו- $S_i\cap S_j=\emptyset$. מקבלים k קבוצות ב"ת שמכסות את כל הקלסטרים, ולכן $\chi(G)\leq k$.

<u>תקפות</u>: אם בגרף המקורי לא יכולנו לצבוע יותר מ-$\delta\cdot n$ צמתים, אז כל קבוצה ב"ת בגרף הקלסטרים היא מגודל $\geq\delta\cdot n$. כיוון ש-$\chi(G)\geq\frac{n}{\alpha(n)}$ אז $\chi(G)\geq\frac{k\cdot n}{\delta\cdot n}=\frac{k}{\delta}$ כאשר $\frac{k}{\delta}$ גדול כרצוננו.

<u>מסקנה</u>: הבעיה $Gap-\chi\left[k,\frac{k}{\delta}\right]$ היא NP-קשה, ולכן קשה לקרב את בעית מספר הצביעה בכל פקטור קבוע.

<u>$Gap-UG[\varepsilon,1-\varepsilon]$</u>:

בעית גרף אילוצים בה לכל קשת (u,v) צביעת קודקוד אחת קובעת את הצביעה של השניה, כאשר המטרה לספק כמה שיותר אילוצים על הקשתות שהן פרמוטציות צבעים. לא ידוע האם בעיה זו ב-P או ב-$NP-hard$.

<u>מציאת הצבה מספקת</u>: צובעים קודקוד כלשהו בכל אחד מרכיבי הקשירות, והוא קובע את צבע שאר הקודקודים באותו רכיב, ממשיכים כך עד שיתקיימו כמה שיותר אילוצים.

בעית $Gap-MaxCut\left[1-\sqrt{\varepsilon},1-\varepsilon\right]$:

מציאת חתך בגרף כך שרק חלק מקשתות הגרף נמצאות לא באחד מצידי החתך ולא בחתך עצמו. האילוץ מתבטא באמצעות שני צבעים אפשריים לכל קודקוד כאשר $\Pr[A(u)\neq A(v)]$ היא הסתברות הקשת (u,v) להימצא מחוץ לחתך. מקרה פרטי של בעית $Gap-UG$ רק עם שני צבעים. ידוע ש-$Gap-MC[1-1.01\varepsilon,1-\varepsilon]\in NP-hard$.

<u>פרק רביעי: חישובים פרובבילוסטיים</u>:

<u>המחלקה BPP</u>: כל הבעיות הניתנות לפתרון פולי' עם שימוש בראנדומיות. **לא ידוע האם $BPP\subseteq NP$**.

<u>הגדרת PH</u>:

נגדיר את Σ_i: תת קבוצה של המחלקה $TQBF$ של כל הנוסחאות מהצורה $\exists x_1\forall x_2\ldots\varphi$ - מתחילות ב-$\exists$ ולאחריהן $i-1$ החלפות כמתים. נגדיר את ההיררכיה הפולינומיאלית: $PH=\bigcup_i\Sigma_i$. נשים לב שבניגוד ל-$TQBF$, מספר ההחלפות לא תלוי בגודל הנוסחה. מחלקה זו סגורה תחת $Karp$.

באופן דומה ניתן להגדיר $\Pi_i=co-\Sigma_i$.

מתקיים: $\Sigma_1=NP,\Pi_1=coNP$, ולכל i: $\Sigma_i,\Pi_i\subseteq\Sigma_{i+1}$ וגם $\Sigma_i,\Pi_i\subseteq\Pi_{i+1}$. כמו כן: $PH\subseteq PSPACE$.קריסת ההיררכיה הפולינומיאלית: אם $NP=coNP$, באינדוקציה על i.

<u>מחלקת PP</u>: $L\in PP$ אם קיימת מכונה פרובבילוסטית פולי' כך ש-$x\in L\Rightarrow\Pr_r[M(x,r)=' x\in L']>\frac{1}{2}$, ו-$x\notin L\Rightarrow\Pr_r[\ldots]\leq\frac{1}{2}$.

<u>מכונת טיורינג פרובביליסטית</u>:

בנוסף לסרט הקלט וסרט העבודה יש סרט רנדומי r, וכדי להחליט האם המכונה מקבלת את x מסתכלים על ההסתברות $\Pr_r[M(x,r)]$.

$L \in BPP$ אמ"מ קיימת M מכונה פרוב'-פולי' כך ש-$\forall x Pr_r[M(x,r) =' x \in L'] > \frac{2}{3}$, ו-$\Pr_r[M(x,r) =' x \notin L'] < \frac{1}{3}$.

טענה: $NP \subseteq PP$:

הפיכת מכונת NP הבודקת האם קיימת ריצה אי-דטר' מקבלת – כלומר מסתכלת על מספר הריצות המקבלות ומבדילה בין 0 ל-> 0 למכונה פרוב':

מוסיפים למכונה צעד התחלתי המגריל מספר בין 0 ל-1. אם 0, דוחים. אם 1, ממשיכים כרגיל. אם קיבלנו 0, במצטבר קיבלנו $\frac{1}{2}$, ודוחים (כי $\geq \frac{1}{2}$).

אם קיבלנו שההסתברות להצלחה היא $\frac{1}{2} + \varepsilon > \frac{1}{2}$, מקבלים.

אמפליפיקציה:

<u>טענה</u>: אם $L \in BPP$ אז קיימת TM פרוב' פולי' M' כך שלכל $x \in \{0,1\}^n$: $\Pr_{r \in \{0,1\}^{p(n)}}[M'(x,r) \neq \chi_L(x)] < \frac{1}{3 \cdot p(n)}$, כאשר χ_L היא פונ' הזהות של השפה (מחזירה 1 אם x בשפה ו-0 אחרת). כלומר: ההסתברות לטעות קטנה מ-$\frac{1}{3 \cdot p(n)}$.

משפט: $BPP \subseteq \Sigma_2^P$ (כאמור לא ידוע $BPP \subseteq NP$):

תהי $L \in BPP$, אז קיימת M כמוגדר לעיל כך ש: $x \in L \Leftrightarrow \exists s_1, \dots, s_m \in \{0,1\}^m . \forall r \in \{0,1\}^m . \bigvee_{i=1}^m M(x, r \oplus s_i)$, והסתברות השגיאה של M היא $\frac{1}{3m}$, ולכן ההסתברות של M לקבל היא $\frac{1}{3m}$. ה-xor משמעותו הזזה של הביטים. נשתמש בתכונה ההסתברותית: $\Pr[a\ has\ property\ p] \Rightarrow exists\ such\ a\ with\ property\ p$.

$x \notin L$: במקרה זה רק $\frac{1}{3m}$ מהממחרוזות r הן טובות, ולכן גם אם נזיז את כולן (ע"י xor עם s_i כלשהו) כך שכולן יהפכו לטובות, תוך שימוש ב-$union\ bound$ (במקרה הטוב ביותר שכל מחרוזת טובה שונה ממחרוזות טובה אחרת שמתקבלת), נקבל $m \cdot \frac{1}{3m} = \frac{1}{3}$ מחרוזות, ולכן M תדחה.

$x \in L$: כל המחרוזות r האפשריות הן $\{0,1\}^m$, ומתוכן $\frac{1}{3m}$ לא טובות. נראה שע"י הזזות עם m מחרוזות s_i נקבל שכל r (מקורית או לאחר הזזה) גורמת ל-M לקבל. נראה שההסתברות לבחירת $s_1, \dots, s_m$ כאלה גדולה מ-0 ולכן לפי התכונה ההסתברותית, קיימים כאלה מחרוזות, ו-M תקבל:

הסתברות המקרה המשלים: ההסתברות שקיים r רע לכל בחירת s-ים: $\leq \Pr_{s_1,\dots,s_m \in \{0,1\}^m}[\exists r \in \{0,1\}^m, \bigwedge_{i=1}^m M(x, r \oplus s_i) = 0]$

נסתכל על כל ה-rים, וזה $overcounting$. אם זה יהיה הסת' קטנה, זה טוב: $\leq \sum_{r \in \{0,1\}^m} \Pr_{s_1,\dots}[\bigwedge_{i=1}^m M(x, r \oplus s_i) = 0] \leq$

המעבר למכפלה מכך שהמאורעות ב"ת, כי r ו-s_i רנד', ולכן גם $r \oplus s_i$: $\leq \sum_{r \in \{0,1\}^m} \prod_{i=1}^m \Pr_{s_1,\dots}[M(x, r \oplus s_i = 0] \leq$

2^m הוא מס' ה-r האפשריים, s_{rand} מחרוזת רנדומית: $\leq 2^m \cdot \prod_{i=1}^m \Pr_{s_{rand} \in \{0,1\}^m}[M(x, s_{rand}) = 0] = 2^m \cdot \left(\frac{1}{3m}\right)^m < 1$

ומכאן ההסתברות המשלימה גדולה מ-0, ולכן קיימים $s_1, \dots, s_m \in \{0,1\}^m$ כך שלכל r תתקיים הנוסחה, ולכן הנוסחה נכונה.

מכאן ש-$L \in \Sigma_2^P$ ולכן $BPP \subseteq \Sigma_2^P$.

***Undirected-CONN* ב-*logspace* ע"י אלג' רנד'**:

נניח הגרף קשיר. נוסיף לכל קודקוד קשת עצמית. נתחיל ב-s כאשר דרגתו ההתחלתית היא $\frac{1}{d_i}$. נבחר בהסתברות $\frac{1}{d_i}$ לכל אחד משכניו שכן אחד, ונמשיך כך. יהי v_t סימון הקודקוד בו נמצאים לאחר t צעדים, $s = v_0$. נגדיר: $P_t(i) = \Pr[v_t = i]$ פונ' התפלגות. למשל: $P_0(s) = 1$ כי הסיכוי של s להיות הקודקוד בו נמצאים בזמן 0 הוא 1. לכל i מתקיים: $\lim_{t \to \infty} P_t(i) = \frac{d_i}{2|E|}$ - התפלגות לינארית בדרגה (בגרף קשיר).

<u>למה</u>: אם ל-t כלשהו מתקיים לכל i: $P_t(i) = \frac{d_i}{|E|}$ אז לכל i: $P_{t+1}(i) = \frac{d_i}{|E|}$. לפיכך, התפלגות כל הקודקודים היא $\frac{d_i}{|E|}$.

לכל קודקוד i, תוחלת מספר הצעדים שנעבור מיציאה ממנו ועד חזרה אליו היא $\frac{2|E|}{d_i}$. אם כן, נבדוק כמה צעדים יקח מ-s ל-t (תוחלת מס' הצעדים):

כדי להגיע מ-s לקודקוד הבא במסלול ל-t יש סיכוי $\frac{1}{d_i}$ ללכת בכיוון הנכון. אם לא, יקח בממוצע $\frac{2|E|}{d_i}$ לחזור ל-s ולבחור בשכן הבא. לכל היותר לאחר d_i נסיונות צפויים לבחור בשכן הנכון, סה"כ: $d_i \cdot \left(\frac{2|E|}{d_i}\right) = 2|E|$ צעדים. לפיכך מסלול מ-s ל-t יקח לכל היותר (הסתברותית) $2|E||V|$.

האלג': נתחיל ב-s ונבחר שכן באופן רנדומי, וכך הלאה עד שנגיע ל-t. אם תוך $2|E||V|$ צעדים לא הגענו ל-t, נדחה.

<u>תרגול</u>:

<u>שיטת התוחלות המותנות</u>:

Max-Cut:

<u>המטרה</u>: למצוא באופן דטר' חתך שגודלו לפחות חצי מקשתות הגרף. בהינתן השמה חלקית של צמתי הגרף לצידי החתך, ניתן לחשב את תוחלת גודל החתך עבור הגדרה רנדומית וב"ת של שאר הצמתים.

<u>הרעיון</u>: נבנה השמה ב-n שלבים. בכל שלב נבחר צד לאחד הצמתים: נחשב את תוחלת גודל בחתך בשלב ה-i לכל בחירה של צד, ובהינתן ההשמה ל-$1,\dots,i-1$ הצמתים הקודמים. שני הערכים שנקבל לא יכולים להיות קטנים מהתוחלת לאחר השלב ה-$i-1$, ולכן תוחלת זו לא קטנה לאורך האלג'. לכן בזמן פולי' מקבלים פתרון שערכו לפחות כמו תוחלת פתרון מקרי.

אלגוריתם רנדומי למציאת *Min-Cut* (*Karger, 94'*):

<u>המטרה</u>: מציאת חלוקת V לשתי קבוצות V_1, V_2 לא ריקות כך שמס' הקשתות ביניהן מינימלי (פתרון רנדומי, לא כבעית זרימה).

<u>אלגוריתם</u>: כל עוד יש יותר משני צמתים בגרף, נקח קשת ונכווץ אותה. קבוצת השכנים החדשה היא איחוד קב' השכנים של הצמתים המקוריים, ומספר הצמתים יורד ב-1. במהלך האלג' נוצרות קשתות כפולות, עליהן שומרים, ולולאות (מצומת לעצמו) זורקים בסיום הריצה. מתקבל גרף עם שני צמתים כאשר האינדקס המצטבר של כל אחד מהם הוא החלוקה של החתך.

<u>רעיון האלג'</u>: בהינתן חתך C, אם במהלך האלג' כיווצנו קשת ששיכת לחתך, לעולם לא יוחזר C בסיום הריצה ("נדפק"). אם בשום שלב לא בחרנו קשת מ-C, בסוף הריצה הוא יוחזר. גודל חתך מינימלי <u>לא יורד</u> במהלך האלג': כל חתך ששורד עד השלב ה-i, מס' הקשתות בו לא יורד. אם גודל חתך מינימלי הוא k, אז כל צומת מדרגה לפחות k (אחרת יש חתך בו צומת מדרגה קטנה מ-k בצד אחד ושאר הקודקודים בשני). אם בגרף n קודקודים אז מס' הקשתות הכולל הוא לפחות $\frac{kn}{2}$.

יהי C חתך מינימלי כלשהו בגרף המקורי. ההסתברות שישרוד את החתך הראשון (ששום קשת בו לא תלקח): $1-\frac{k}{|E|} \geq 1-\frac{k}{kn/2} = 1-\frac{2}{n}$. בשלב השני יש לפחות $\frac{k(n-1)}{2}$ ($n-1$ צמתים, גודל חתך מינימלי לא יורד) קשתות בגרף, וההסתברות לשרוד שלב שני$\leq 1-\frac{2}{n-1}$. באופן כללי הסיכוי להשרדות בשלב ה-$i \leq \frac{n-i-1}{n-i+1} = 1-\frac{2}{n-i+1}$. ההסתברות שהחתך ישרוד עד הסוף הוא טור טלסקופי המצטמצם לכדי $1/\binom{n}{2} = \frac{2}{n(n-1)}$.

חזרה על ריצת האלג' $O(n^2 \log n)$ פעמים תתן את החתך המינימלי בהסתברות גבוהה.

<u>מסקנה קומבינטורית</u>: כל חתך מינימלי מתקבל באלג' בהסתברות לפחות $1/\binom{n}{2}$. בנוסף, המאורעות בהם בחרנו חתכים שונים זרים. מכאן, מספר החתכים המינימלי בגרף $\geq \binom{n}{2}$, וזה חסם הדוק.

<u>חסם צ'רנוב</u>:

משפט צ'רנוב (חסם צ'רנוב):

יהיו $X_1,\dots,X_k \in \{0,1\}$ משתנים מקריים ב"ת, $X = \frac{1}{k}\sum_{i=1}^{k} X_i$ (תוחלת). אז הסיכוי לטעות: $\Pr\left[\left|\frac{1}{k}\sum_{i=1}^{k} X_i - E[x]\right| > \delta\right] \leq 2 \cdot e^{-2\delta^2 \cdot k} = \varepsilon$.

מתקיים: $\ln\left(\frac{1}{\varepsilon}\right) = \ln\left(2 \cdot e^{-2\delta^2 k}\right) \Leftrightarrow 2 \cdot e^{-2\delta^2 k} = \varepsilon$ ולכן $k = O\left(\ln\frac{1}{\varepsilon} \cdot \frac{1}{\delta^2}\right)$. מכאן ש-$\delta$ תלוי רק בגודל הודאות הנדרשת ולא בגודל המדגם.

אלג' קירוב לבעית *Set-Cover* תוך שימוש ב-*L.P.*:

נעביר את הבעיה לבעית LP:

- לכל קבוצה $S_i \in F$ נתאים משתנה x_i ונקבל $\{x_1,\dots,x_k\}$.
- אם $s_i \in C$ (הכיסוי), אז $x_i = 1$, אחרת $x_i = 0$.
- רוצים להביא למינימום את $min\sum_{i=1}^{k} x_i$ כך ש-$\forall j \in \{1,\dots,n\}. \sum_{i:j\in S_i} x_i \geq 1$ (כלומר: לפחות קבוצה אחת ב-C מכילה את j, לכל $j \in \{1,\dots,n\}$).
- $x_i \in \{0,1\}$.

הבעיה לעיל היא IP ולכן היא NP-קשה, נעביר לבעיה מקורבת שהיא LP ע"י הפיכת האילוץ האחרון ל-$x_i \in [0,1]$. פולי' עם אלג' האליפסואיד.

יהי P פתרון אופטימלי לבעית ה-LP (שקיבלנו): $P =< p_1,\dots,p_k >$. ידוע כי $OPT = \sum_{i=1}^{k} p_i$. $OPT \leq set-cover$. אלג' הקירוב ישתמש בעיגול התוצאה כדי למצוא SC חוקית. מהפתרון נוציא SC עם t קבוצות: $set-cover\ with\ t\ sets \leq t \leq O(\log n) \cdot set-cover$.

<u>האלג' הרנדומי</u>: משתמש ב-p_i מפתרון ה-LP כהסתברויות לבחירת S_i שיכנסו לפתרון C.

נקח כל קבוצה S_i בהסתברות p_i אל הפתרון C. תוחלת: $E[\#sets\ chosen] = \sum_{(i=1)}^{k} p_i = OPT$. כלומר, תוחלת הפתרון לפי אלג' זה זהה ל-OPT. נחזור על פעולה זו $d \cdot \log n$ פעמים כאשר d יקבע אח"כ. כדי להבטיח כיסוי מושלם לכל j.

מתקיים: $E[\#sets\ in\ C] \le d \cdot \log n \cdot OPT$. נניח j נמצא ב-f קבוצות מתוך F: $\Pr[j\ \textbf{not}\ covered\ in\ some\ round] = \prod_{l=1}^{f} p_{i_l} \le \prod_{l=1}^{f} e^{-p_{i_l}} = e^{-\sum_{l=1}^{f} p_{i_l}} = e^{-\sum_{i_l : j \in S_{i_l}} p_{i_l}} \le \frac{1}{e}$, כלומר לבסוף קיבלנו חסם קבוע.

נבצע $d \cdot \log n$ איטרציות ובסיכוי גבוה נקבל שכל j מכוסים: $\forall j.\ \Pr[j\ covered\ in\ some\ round] \ge 1 - \frac{1}{e} (\cong 60\%)$. כדי שכל j יכוסו נבחר d כך ש-$\left(\frac{1}{e}\right)^{d \cdot \log n} \le \frac{1}{4n}$. כל איטרציה ב"ת בקודמתה ובכל איטרציה הסיכוי לכל j שלא יכוסה הוא $\frac{1}{e}$. הסיכוי שלא הצלחנו לתפוס את j בכל האיטרציות: $\left(\frac{1}{e}\right)^{d \cdot \log n} = \frac{1}{4n}$. יש n j-ים שונים, ולכן יש *union-bound*: $\forall j.\ \Pr[j\ not\ covered\ in\ all\ rounds] \le \frac{1}{4n} \cdot n = \frac{1}{4}$. מכאן הסתברות של $\frac{3}{4}$ לקבל SC טוב, שהוא C.

SDP (*Semi-Definite Prog.*):

מוגדרת כמו בעית LP אבל: $v_1, \dots, v_n \in \mathbb{R}^n$ וקטורים. תזכורת: $v = (v^1, \dots, v^n), w = (w^1, \dots, w^n)$, $vw = \sum_{i=1}^{n} v^i w^i$, $\|V\| = \sqrt{v \cdot v}$.

$(*) = \min \sum_{i,j=1}^{n} A_{i,j} \cdot v_i \cdot v_j$ כאשר $v_1, \dots, v_n \in \mathbb{R}^n$ אילוצים הם מהצורה $\sum_{i,j=1}^{n} B_{i,j}^{k} \cdot v_i \cdot v_j \ge b^k$ עבור $k \in \{1, \dots, m\}$. ה-SDP מוצא וקטורים $v_1, \dots, v_n \in \mathbb{R}^n$ המביאים למינימום את $(*)$.

ל-SDP אלג' יעילים הפותרים אותו בזמן פולי'.

Max-Cut עם SDP:

בהינתן $G = (V, E)$ לא מכוון רוצים למצוא $S \subseteq V$ כך שגודל קב' הקשתות (u, v) כך ש-$u \in S, v \notin S$ מקסימלית. בעיה זו NPC. להלן אלג' קירוב המשתמש ב-SDP ונותן פקטור 0.878, כלומר נותן פתרון S' כך ש-$|E(S', \bar{S'})| \ge 0.878 \cdot OPT$ ($OPT = |E(S, \bar{S})|$). נכתוב כ-IP:

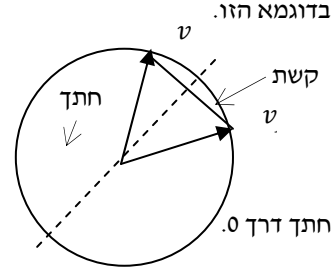
לכל $i \in V$ נתאים x_i כך ש-$x_i = \begin{cases} 1, i \in S \\ -1, i \notin S \end{cases}$. מכאן שמתקיים לכל $i, j \in V$: אם $x_i \cdot x_j = 1$ אז i, j נמצאים או לא נמצאים ב-S ביחד, כלומר הקשת (i, j) לא בחתך. באופן דומה, אם מכפלת $x_i \cdot x_j = -1$ אז הקשת (i, j) בחתך.

אם נקח $\frac{1 - x_i \cdot x_j}{2}$ נקבל: $\begin{cases} 1, (i,j) \in CUT \\ 0, (i,j) \notin CUT \end{cases}$. הבעיה שנרצה למצוא: $\max \sum_{(i,j) \in E} \frac{1 - x_i \cdot x_j}{2}$ כך ש-$\forall i \in V.\ x_i \in \{-1, 1\}$.

<u>הפיכת הבעיה לבעית SDP</u>:

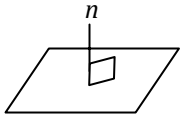
וקטורים $v_1, \dots, v_n$, הבעיה: $\max \sum_{(i,j) \in E} \frac{(1 - v_i \cdot v_j)}{2}$ תחת האילוץ: $\forall i \in V.\ v_i \cdot v_i = \|v_i\| = 1$, כלומר לכל צומת מתאים וקטור יחיד. שאיפה שכולם יהיו הפוכים אחד לשני כדי שמכפלתם תהיה -1, אך זה כמובן לא אפשרי. בדוגמא עבור 3 וקטורים:

$\max \frac{1}{2}(1 - v_1 v_2 + 1 - v_2 v_3 + 1 - v_3 v_1) = \frac{3}{2} - \min \frac{1}{2}(v_1 v_2 + v_2 v_3 + v_3 v_1)$, כאשר עבור 3 וקטורים הפתרון האופטימלי יהיה אם ביניהם $120°$, כלומר $v_i \cdot v_j = \cos 120 = -\frac{1}{2}$, ולכן הנ"ל שווה ל-$2\frac{1}{4} \ge \frac{3}{2} - \frac{\frac{1}{2}+\frac{1}{2}+\frac{1}{2}}{2}$, כאשר 2 הוא ה-*mincut* בדוגמא הזו.



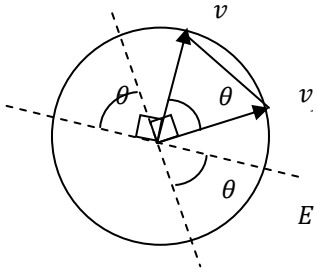
<u>אלג' *Goemans-Williamson* (94')</u>:

- תחילה נפתור את בעית ה-SDP ונקבל $v_1, \dots, v_n$ כאשר $OPT - SDP = \sum_{(i,j) \in E} \frac{1 - v_i v_j}{2}$.
  לכל צומת יש וקטור, אותו נשים על מעגל (יותר נכון ספירה, שכן במימד n).
- <u>האלג'</u>: מעבירים חתך רנדומי, וכך הסיכוי הגדול ביותר לחתוך מקסימום מהקשתות. בשני מימדים – מעבירים חתך דרך 0.
  ב-n מימדים מעבירים *hyper-plain* – מישור ממימד $n - 1$ במרחב ממימד n.



תיאור מישור רנדומי: כל מישור רנדומי מתואר ע"י וקטור נורמלי (יחיד) n (המישור עובר דרך 0). תיאור: $plain = \{v |\ n \cdot v = 0\}$. מציאת מישור רנדומי שקולה למציאת הוקטור הנורמלי שלו n רנדומי, $\|n\| = 1$. בוחרים וקטור כלשהו מהספירה, והוא יהיה n.

כעת נבדוק כמה קשתות נחתכו ע"י המישור לאחר בחירת n:



$$\Pr[(i,j) \in E\ is\ cut\ by\ plain] = \frac{2\theta}{2\pi} = \frac{\theta}{\pi}$$

<u>טענה</u>: $\Pr[\dots] \ge 0.878 \cdot \frac{1 - v_i v_j}{2}$.

אם הטענה נכונה אז $E[CUT] = \sum_{(i,j) \in E} \Pr[(i,j) \in E\ is\ cut] \ge 0.878 \cdot \sum_{(i,j) \in E} \frac{1 - v_i v_j}{2} \ge 0.878 \cdot MAX - CUT$

<u>הוכחת הטענה</u>:

$$\frac{\Pr[(i,j) \in E\ is\ cut]}{\frac{1 - v_i v_j}{2}} = \frac{\frac{\theta}{\pi}}{\frac{1 - v_i v_j}{2}} = \frac{\frac{\theta}{\pi}}{\frac{1 - \cos\theta}{2}} = \frac{2}{\pi} \cdot \frac{\theta}{1 - \cos\theta} \underset{\text{מבדיקת מינימום}}{\ge} 0.878$$