קושי של קירובים

עפ"י הרצאות של ד"ר גיא קינדלר

סמסטר א', תש"ע

סיכם: שיר פלד, באמצעות LyX גרסה 1.6.1
תיקונים יתקבלו בברכה בכתובת מייל $shirpeled@cs$

1 מבוא היסטורי

ישנם שני נושאים שהתפתחו ולבסוף נפגשו ויצרו את הנושא הנ"ל:

קושי של בעיות אופטימיזציה (וקירובן)

בעיית אופטימיזציה היא אוסף של קלטים שמיוצגת ע"י X, כך שהתשובה היא S (מחרוזת) כלשהי. בעית האופטימיזציה מיוצגת ע"י הצורך למקסם פונקציה כלשהי $f(X,S) \to \mathbb{R}^+ \cup \{\perp\}$, כך שלכל פתרון נותנים משקל או סימן מיוחד המייצג את העובדה ש S כלל אינה פתרון עבור X.
אפשר כמובן גם לבחור בבעיית מינימיזציה, לדוגמה: צביעה של גרף. X יהיה גרף. S תהיה צביעה של הגרף, כך ש $f(X,S)$=מספר הצבעים ש S משתמשת בהם (אם היא צביעה חוקית ו$\perp$ אחרת).
בעיה נוספת היא $Exact\,3-CNF-SAT$, כאשר הקלט הוא $(x_1 \vee x_2 \vee x_3) \wedge (\bar{x_1} \vee x_2 \vee \bar{x_3}) \wedge \ldots$.
מרחב הפתרונות הוא כל ההשמות הבוליאניות למשתנים. אם נסמן השמה כלשהי ב I, אזי נגדיר את פונקציית המשקל להיות מספר הסוגריים שערכם ביחס להשמה $f(I,S)$ הוא אמת.
בעיות נוספות: כיסוי קודקודים, כיסוי קבוצות, קבוצה בת"ל ועוד...
אפשר לגזור מכל בעיית אופטימיזציה בעיית הכרעה. למשל - עבור קבוצה בת"ל - נבקש עבור גרף נתון לענות על השאלה האם עבור k נתון יש קבוצה בלתי תלויה בגודל k. מה שמעניין הוא גם שברוב המקרים זה גם עובד בכיוון השני. בבעיה הקודמת למשל - נשאל את שאלת ההכרעה עבור $k=1,2,3,\ldots$. שאלה נוספת היא איך ניתן למצוא קבוצה בלתי תלויה מסויימת בגודל k - וגם לכך יש אלגוריתם פשוט (היה בקורס אלגוריתמים).

הגדרה 1.1 קירוב α לבעיית מקסימיזציה $f(,)$: אלגוריתם אשר בהנתן קלט I מוצא S כך ש- $f(I,S) \geq \alpha \cdot \max_S f(I,S)$, וקירוב α לבעית מינימיזציה מוגדר ע"י $f(I,S) \leq \frac{1}{\alpha} \cdot \min_S f(I,S)$.

מאוחר יותר הוגדרה המחלקה NP ע"י $Cook$ ו $Levin$, והם הראו שקיימת בעיה שלמה ב NP והיא : $E3-CNF-SAT$. עד היום רוב בעיות האופטימיזציה נמצאות ב NP. התוצאה היתה די אזוטרית, עד שהוכיח $Karp$ שיש אוסף אדיר של בעיות שגם הן NP שלמות: כיסוי קבוצות, כיסוי קודקודים, חיתוך מקסימלי בגרף, צביעה, ועוד...
עד תחילת השמונים עיקר העיסוק היה לנסות להוכיח ש $P \neq NP$ ומאידך להוכיח שבעיות מסוימות הן NP קשות. הנסיון להוכיח ש $P \neq NP$ לא נשא פרי עד היום, וגם הרדוקציות הפכו להיות די סטנדרטיות. החל משנות השמונים הכיוון היה למצוא קירובים עבור בעיות NP-קשות. אחת התוצאות הראשונות (מ 76) בענף הקירובים היא שדי קל למצוא 2 - קירוב לבעיית $Max-cut$ (למצוא חתך מקסימלי בגרף ממושקל לא-מכוון). בשנת 94 מצאו $Goemans$ ו $Williamson$ מצאו קירוב $0.878\ldots$ לבעיית החיתוך המקסימלי ע"י שיכון הגרף ב $\mathbb{R}^n$ ומציאת פתרון גיאומטרי, ועד היום לא הצליחו לשפר את הקירוב הזה ואף לא להגיע לאותו קירוב באמצעים שאינם גיאומטריים.
נזכיר את בעיית ה $E3-CNF-SAT$, שאותה ניתן לקרב ע"י השמה רנדומית. מכיוון שעבור שלושה ליטרלים המשורשרים ע"י $\vee$ יש רק השמה אחת שנותנת $False$, בתוחלת נקבל קירוב של $\frac{7}{8}$ (יש 8 השמות אפשריות). יש עם זה כמה בעיות - השימוש ברנדומיות ומאידך ההסתמכות על תוחלת.
אם משלבים את הרעיון של חלוקה למחלקות סיבוכיות עם הרעיון של קירובי בעיות אופטימיזציה אפשר לשאול - אולי ניתן להראות שקירוב של יותר מ $\frac{7}{8}$ לבעיה האחרונה הוא בעיה NP קשה? מסתבר שהטכניקות שהיו לנו קודם לכן - לא טובות לשם כך.

$PCP - Theorem$ (הוכחות ניתנות לווידוא הסתברותי, הוכח בשנת 92)

הגיע מתחום אחר לגמרי של עיסוק - של העברת מידע והצפנה ומתישהו בשנות התשעים הסתבר שיש קשר בין הדברים.
$Cook$ הבחין בכך (במאמרו המקורי על בעיות NP) שבעיות NP מגלמות בתוכן הוכחות לטענות מתמטיות שניתן לוודא בזמן פולינומי. נניח ש I הטענה ויש מקום של n ביטים להוכחה. $Cook$ הראה שניתן לתרגם את הטענה למערכת $n - CNF - SAT$ שהיא ספיקה אם"ם הטענה נכונה.

הגדרה 1.2 IP - פרוטוקול אינטראקטיבי, שבו ה $Verifier$ מנסה לוודא טענה של ה $Prover$. למשל - המוכיח מנסה להראות שיש לו אסטרטגיה מנצחת עבור משחק כלשהו. נרצה שהפרוטוקול יהיה כזה שבסיום הבדיקה, אם יש אסטרטגיה מנצחת אז בהסתברות $\frac{2}{3}$ המוודא יווכח בכך, ואם אין - אז בהסתברות של $\frac{2}{3}$ המוודא יעלה על השקר. מסתבר שאם מוסיפים ל $Verifier$ גם כוח של רנדומיות - זה ניתן.
ZK - הוכחות אפס מידע. הרעיון הוא דומה ל IP אלא שהמוודא ישתכנע בוודאות גבוהה אם יש אסטרטגיה מנצחת, אבל הוא לא יקבל מכך שום מידע מהי האסטרטגיה המנצחת הזו.
AM - ארתור ומרילין במקום מוודא ומוכיח, וכל השאר דומה מלבד העובדה שמקור הרנדומיות הוא משותף לארתור ומרילין, ומסתבר שזה בדיוק שקול ל IP .

בהמשך הוכח (ע"י שמיר) ש $IP = PSPACE$ ומאידך (הראו אחרים) כי $\exists A : Co - NP^A \not\subseteq IP^A$ (כאן העלאה בחזקה משמעו הנחת אורקל לשפה A). ואחריו הראו ש $MIP = NEXP$ (MIP היא המחלקה שבה יש מספר $Provers$ שאינם יכולים לתאם ביניהם ולכן זה מוסיף כוח כי המוודא יכול לשאול אותם ולהשוות).
נשאלה השאלה האם אפשר להגביל את הפרוטוקולים ב MIP ולקבל משהו ששקול ל NP? למה זה יתכן? כי אם בזמן פולינומי אפשר בהסתברות $\frac{2}{3}$ לוודא בעיה ענקית שהיא ב $NEXP$ - אז אולי אפשר באמצעים דומים לקבל קירוב הסתברותי לווידוא של בעיה ב NP, שהיא יותר פשוטה מן הסתם.

הגדרה 1.3 $PCP\left[r\left(n\right),q\left(n\right)\right]$
נתונה הוכחה באורך n. פרוטוקול לבדיקת נכונות ההוכחה הוא ב $PCP\left[r\left(n\right),q\left(n\right)\right]$ אם מתקיימים התנאים הבאים:

1. רץ בזמן פולינומי ב n
2. משתמש לכל היותר ב $r\left(n\right)$ הגרלות של ביטים רנדומיים
3. קורא לכל היותר $q\left(n\right)$ ביטים מתוך ההוכחה
4. אם הטענה נכונה וגם ההוכחה נכונה אזי

$$Pr\left(Protocol\,accepts\right)=1$$

5. אם הטענה אינה נכונה אזי

$$Pr\left(Protocol\,accepts\right)\leq\frac{1}{3}$$

ב 91 הוכיחו (פייגה, גולדרייך, לובאץ' וספרא) כי לכל שפה ב NP יש:

$$f\left(n\right)=\log\left(n\right)\cdot\log\log\left(n\right)$$

כך שהשפה שייכת ל $PCP\left[f\left(n\right),f\left(n\right)\right]$.
ומכאן הם הראו שנובע שלכל קבוע c, אם יש אלגוריתם המבטיח c-קירוב לבעיית ה $Max-Clique$ בזמן פולינומי, אזי $NP\subseteq TIME\left[n^{\log\log(n)}\right]$. בדיעבד - זה המאמר שגרר את ההתפתחות של תחום המחקר של קושי חישובי.

משפט 1.4 $NP=PCP\left(\log\left(n\right),1\right)$.
כלומר יש הוכחה וטענה ששתיהן יחד באורך m - המוכיח מתרגם את ההוכחה למשהו באורך n בפורמט מוסכם, כך שאורך התרגום הוא לא יותר מפולינומי ב m. אפשר לקרוא מספר קבוע של ביטים (תוך הגרלת $\log\left(n\right)$ ביטים) ולקבוע בוודאות מסויימת האם ההוכחה נכונה או לא. יש שיפורים כך שהאורך של n יהיה בסדר גודל של $m\cdot poly\left(\log\left(m\right)\right)$.

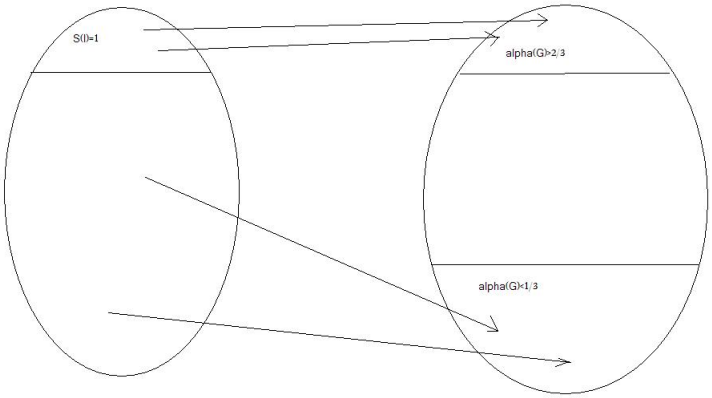
2 קושי של בעיות קירוב ו PCP

הגדרה 2.1 בעיית אופטימיזציה : $f(I,S) \to \mathbb{R}^+ \cup \{\perp\}$, בהנתן I מצא S כך ש f מקסימלית.

קירוב α - מצא S כך ש- $f(I,S) \geq \alpha \cdot \max_S f(I,S)$

הגדרה 2.2 המחלקה NPO היא מחלקת בעיות האופטימיזציה עבורן כל פתרון פיזיבילי S הוא באורך פולינומי ב $|I|$, וכן $f(I,S)$ ניתנת לחישוב בזמן פולינומי.

מעכשיו כשנאמר "בעיית אופטימיזציה" נתכוון לבעיה ב NPO.
איך נראה שבעיה היא NPO קשה? כלומר קשה לקירוב?



נסתכל על בעיית הקבוצה הבלתי תלויה בגרף. נגדיר לכל גרף G את $\alpha(G) = \frac{\max IS(G)}{|V(G)|}$ וננסה לסדר את כל הגרפים בעולם בתוך עיגול, שבשלישו העליון נמצאים כל הגרפים המקיימים $\alpha(G) \geq \frac{2}{3}$ ובשלישו התחתון $\alpha(G) \leq \frac{1}{3}$.
באופן דומה נתבונן בבעיית ה $E3SAT$, ונגדיר את $S(I)$ להיות מספר ההסגרים הספיקים המקסימלי לחלק למספר ההסגרים, כלומר החלק היחסי של ההסגרים הספיקים ע"י השמה מיטבית. שוב נפזר בתוך עיגול את כל הבעיות $E3SAT$, כאשר בחלק העליון נשים את $S(I) = 1$ ומתחתיו כל השאר. אנחנו יודעים שלהכריע האם בעיה מסויימת ב $E3SAT$ היא בחלק העליון או בשאר העיגול זה $NP - Hard$.
כעת היינו רוצים רדוקציה פולינומית שתהיה לה התכונה הבאה - את החלק העליון של $E3SAT$ נמפה לשליש העליון של העיגול של הגרפים, ואת שאר העיגול - לשליש התחתון של עיגול הגרפים. אם היתה לנו רדוקציה כזו - היינו מקבלים שבעיית IS הוא $NP - Hard$ לקירוב עד כדי פקטור 2.

הגדרה 2.3 בעיית הפער $\left(\frac{2}{3}, \frac{1}{3}\right) - gap\,IS$: יש להכריע בין המקרים $\alpha(G) \geq \frac{2}{3}$ ו $\alpha(G) \leq \frac{1}{3}$.

.

הגדרה 2.4 בעיית פער (α, β) היא NP-שלמה אם בעיית האופטימיזציה המתאימה היא ב NPO ויש רדוקציה מ $3SAT$ אליה כך ש:

1. הרדוקציה R רצה בזמן פולינומי
2. עבור קלט ספיק I מתקיים $OPT(R(I)) \geq \alpha$
3. עבור קלט לא-ספיק I מתקיים $OPT(R(I)) \leq \beta$

הערה 2.5 אם f היא בעיית אופטימיזציה ב NPO ו $(\alpha, \beta) - gap\,f$ היא NP-שלמה אז זה NP-קשה לקרב את f כדי פקטור $\frac{\beta}{\alpha}$.

נפעיל עיקרון דומה לגבי בעיות צביעה. ידוע שההכרעה האם גרף הוא 3 - צביע היא קשה, ולכן נוכל לבנות עיגול (או אליפסה כמו קודם) שבחלקה העליון יהיו הגרפים שהם 3 -צביעים או פחות מכך, ובחלקה התחתון הגרפים שהם לפחות 4-צביעים. מכאן נקבל ש $(3,4) - gap\,coloring$ היא NP-שלמה.

הערה 2.6 כמו כן $\left(1,\frac{1}{m}\right)-gap3SAT$ היא NP שלמה (כאשר m מספר ההסגרים) אבל זה טיפשי כמובן, כי זו דרך אחרת להגיד ספיק או לא ספיק.
$\left(n,10^6n\right)-gapTSP$ (סוכן נוסע) היא NP קשה. נשנה קצת את ההגדרה ל"צריך לבקר בכל קודקוד פעם אחת בדיוק במחיר מינימלי". נעשה את הרדוקציה המוכרת מ $3SAT$ לבעיית מציאת המילטוניאן בגרף, וניתן לכל הקשתות שמשתתפות בגרף משקל 1 ולכל הקשתות האחרות מחיר 10^6n.
נשים לב שהרדוקציות שאנחנו מכירים עבור בעיות הכרעה מסתמכות על העתקות מקומיות (ליטרלים ב $3SAT$ הופכים לקודקודים בבעיית IS) ולכן אם יש רק ביט אחד שמקלקל (למשל הסגר יחיד שלא מסתפק בשום השמה) - יש רק ביט קטן בצד השני של הרדוקציה (רק שלושה קודקודים) שמשקף את זה, ולכן אותן שיטות לא עובדות עבור רדוקציות של בעיות פער.

נניח שבידינו בעיית $E3SAT$ בניסוח הפער $(1,1-\varepsilon)$, אז נפעיל רדוקציה מוכרת ל IS (ליטרלים עוברים לקודקודים וכו') ונקבל $\left(\frac{1}{3},\frac{1}{3}\left(1-\varepsilon\right)\right)-gapIS$. עבור $\varepsilon=\frac{1}{n}$ זה קל ומקבלים חסם קצת טיפשי, אבל בהמשך זה יהיה שימושי.
נזכיר בעיית $VC\left(G\right)$ - כיסוי קודקודים בגרף, גודל כיסוי הקודקודים המינימלי כך שכל צלע מכוסה לפחות ע"י אחד מקודקודיה. נזכיר שכיסוי קודקודים ו IS הן בעיות משלימות, ולכן כמובן יש רדוקציה מ $E3SAT$ והפער המתאים יהיה $\left(\frac{2}{3},\frac{2}{3}+\frac{\varepsilon}{3}\right)$.

הערה 2.7 בעיית אופטימיזציה תלויה בפונקציה שנותנת משקל לפתרונות. לאותה בעיה אפשר לתת פונקציות שונות ולכן צריך לשים לב שההגדרה של הבעיה היא טבעית וטובה.

הגדרה 2.8 פרוטוקול $PCP^{na}\left[r\left(n\right),q\left(n\right)\right]$ עבור שפה L, הוא אלגוריתם אשר מקבל קלט I באורך n, מחרוזת אקראית באורך $r\left(n\right)$, וגישה לפונקציית אוב Δ (על שם בעלת האוב מעין דור).

1. בוחר $q\left(n\right)$ שאילתות לפונקציית האוב
2. מקבל לכל היותר $q\left(n\right)$ תשובות (ביט לכל תשובה)
3. עפ"י התשובות הוא מקבל או דוחה.

בנוסף:

- אם $I\in L$ אז קיימת Δ כך ש $Pr\left[A\left(I\right)=accept\right]=1$
- אם $I\neq L$ אז לכל Δ מתקיים $Pr\left[A\left(I\right)=accept\right]\leq\frac{1}{3}$

הערה 2.9 na בשם הפרוטוקול משמעה שהוא אינו סתגלני, לכן קודם בחירת השאילתות ואז קבלת התשובות ולא שליחת שאילתה ואז קבלת החלטה לגבי המשך השאילתות על סמך התשובה.

הגדרה 2.10 $PCP^{na}\left[r,q\right]$ = מחלקת כל השפות שיש עבורן פרוטוקול כנ"ל.

משפט 2.11 BFL:

$$NEXP=PCP^{na}\left[poly\left(n\right),poly\left(n\right)\right]$$

משפט 2.12 PCP:

$$NP\subseteq PCP^{na}\left[\log\left(n\right),1\right]$$

בהנתן קלט זה אומר שאם מספר הביטים המוגרלים הוא $\log\left(n\right)$ אז יש בעצם $poly\left(n\right)$ התנהגויות אפשריות לאלגוריתם.

טענה 2.13

$$PCP^{na}\left[\log\left(n\right),1\right]\subseteq PCP^{na}\left[\log n,poly\left\{n\right\}\right]\subseteq NP$$

.

טענה 2.14 יש ε כך ש $\left(1,1-\varepsilon\right)-gap\,E3SAT$ היא NP-שלמה.

הוכחה: לטענה הראשונה - ההכלה הראשונה פשוטה. מדוע זה מוכל ב NP? מכיוון שרצף הביטים המוגרל קובע באופן יחיד את המשך הפרוטוקול - אפשר פשוט להריץ את הפרוטוקול על כל הרצפים האפשריים של ביטים (זה פולינומיאלי) , ואם I הוא השמה מספקת - נקבל בכולם שהאלגוריתם יקבל, ואחרת - נקבל שהאלגוריתם דחה בחלק מהפעמים. ■

טענה 2.15 תרגיל לבית: בהנחה ש $P\neq NP$, להראות ש $NP\nsubseteq PCP^{na}\left[r,1\right]$ אם $r\left(n\right)=o\left(\log\left(n\right)\right)$, כלומר הלוגריתם במשפט ה PCP הוא הדוק.

הגדרה 2.16 עבור קבוע q, בעיית האופטימיזציה $q-csp$ מוגדרת כך:

- הקלט I הוא אוסף פרדיקטים בוליאנים על q משתנים בוליאנים.
- פונקציית המטרה: בהנתן קלט והשמה למשתנים, הפונקציה מחזירה את החלק היחסי של הפרדיקטים המסופקים.

משפט 2.17 $q-csp$ כבעיית פער היא NP קשה (עבור q כלשהו).

הוכחה: במקום להסתמך על פונקציית אוב, נאמר שכל הפלטים שהפונקציה יכולה לתת הם $X_1,X_2,...,X_{poly(n)}$, כל אחד באורך ביט, מדוע? כי יש $poly\left(n\right)$ פרמוטציות אפשריות למחרוזת האקראית(=השאלה המופנית לפונקציית האוב), ופונקציית האוב היא חד ערכית. בהנתן I ומחרוזת אקראית באורך $O\left(1\right)=q$, הפרוטוקול יבחר לפי המחרוזת אילו מהביטים של פונקציית האוב הוא בודק, ועליהם נגדיר פרדיקט בוליאני $P\left(X_{k_1},X_{k_2},...,X_{k_q}\right)$ אשר נותן אמת אם הפרוטוקול מקבל. נקבל שאנחנו מייצרים $poly\left(n\right)$ פסוקיות שבכל אחת q ליטרלים. לפי תכונות PCPמובטח שאם I ספיקה אז $q-csp$ ניתנת לסיפוק מלא, ואחרת - בפחות משליש מהפסוקיות, ולכן קיבלנו בעיית פער על $q-csp$ שמהווה את הרדוקציה הדרוש. ■

3 המשך PCP

הערה 3.1 קיבלנו את המשפט הקודם כתוצאה של משפט ה PCP, אבל אם היה נתון לנו המשפט הקודם - ניתן להסיק את משפט ה PCP. מדוע? נניח שהמשפט הקודם נכון, ולכן לכל $L\in NP$, יש רדוקציה של L לבעיית הפער $q-csp\left(1,\frac{1}{3}\right)$. נראה ש $L\in PCP\left[\log\left(n\right),1\right]$. תהי I בעיית הכרעה (כלומר נוסחה ספציפית), אז $R\left(I\right)$ היא הבעיה ב $q-csp$ אחרי הפעלת הרדוקציה. נבקש שהאורקל שלנו יענה על שאלות שהן "מהי ההשמה למשתנה ה-i בנוסחה $R\left(I\right)$". הפרוטוקול יהיה כדלהלן:

1. בחר הסגר אקראי.
2. שאל מה ערכי המשתנים בהסגר הזה.
3. נקבל אם"ם ההסגר מסתפק.

אם $I\in L$ אז $R\left(I\right)$ ספיקה ולכל הסגר שנבחר - ההסגר ספיק והשמתו תינתן על ידי האורקל, ולכן נקבל.
אם $I\notin L$ אז $R\left(I\right)$ לכל היותר $\frac{1}{3}$-ספיקה, ולכן אם נבחר הסגר באקראי, "ניפול" על הסגר ספיק בהסתברות $\frac{1}{3}$, וזה המקרה היחיד שבו נטעה, ולכן נקבל בהסתברות קטנה מ $\frac{1}{3}$.

משפט 3.2 לכל $s>0$ מתקיים שבעיית הפער $q-csp\left(1,s\right)$ היא NP-שלמה (עבור q קבוע שאולי תלוי ב s).

הגדרה 3.3 $PCP^{na}_{c,s}[r,q]$ כאשר c היא פרמטר השלמות $(completeness)$ ונאותות $(soundness)$. כלומר עד כה דיברנו על $PCP^{na}_{1,\frac{1}{3}}[r,q]$.

למה 3.4 $Hardness\,amplification$ - $PCP_{c,s}[r,q] \subseteq PCP_{c^2,s^2}[2r,2q]$

הוכחה: מפעילים את הפרוטוקול פעמיים ונקבל אם"ם הפרוטוקול קיבל בשתי הפעמים, ומכיוון שהמאורעות בלתי תלויים - ההסתברויות נכפלות. ■

זה כמובן נכון באופן כללי, ולכן אם התחלנו עם $c=1$, אז ניתן לחזור על הפרוטוקול מספר פעמים גדול כרצוננו (אך קבוע) וכך לקבל s קטן כרצוננו מבלי לגרוע מ c.

מכאן נובעת הוכחת המשפט - בהוכחת המשפט המקורי הוכחנו עבור שלמות 1 ונאותות $\frac{1}{3}$, כאשר הסתמכנו על q שאילתות לאורקל (כאן זהו אותו q מה $q-csp$) ולכן הקטנת הנאותות תבוא במחיר של הגדלת q.

משפט 3.5 קיים ε קבוע כך שבעיית הפער $(1,1-\varepsilon)\,E3-CNF-SAT$ היא NP שלמה.

הוכחה: אנחנו רוצים רדוקציה שלוקחת בעיות מ $q-csp$ שהן ספיקות לבעיות ספיקות ב $3-CNF-SAT$, ובעיות שהן פחות מ $\frac{1}{3}$ ספיקות - ממירה לבעיות שהן פחות מ $1-\varepsilon$ ספיקות. ננסה להמיר בעיה מ $q-csp$ לבעיית $3-CNF-SAT$ בדרך הנאיבית ביותר ונקווה שזה יעזור. למשל ניקח את הפרדיקט ה i (שפועל על q משתנים): $P_i\left(x_{i_1},...,x_{i_q}\right)$ ונמיר אותו, ע"י פונקציה G לנוסחת $3-CNF-SAT$. מה צריכה G לקיים?

ראשית כל ש $1 \leq |G(P_i)| \leq c$ עבור c קבוע כלשהו.

שנית - אם השמה A מספקת את P_i אז קיימת השמה מספקת $\widetilde{A}$ ל $G(P_i)$, שהיא הרחבה של A. כלומר - יכול להיות שבתהליך המעבר יצרנו משתנים חדשים, אבל אנחנו רוצים שהשמה מספקת ל P_i תיתרגם להשמה מספקת ל $G(P_i)$ שמסכימה ממש עם P_i על כל המשתנים ב P_i.

שלישית - אם A אינה מספקת את P_i, אז כל הרחבה של A להשמה ל $\widetilde{A}$ למשתני $G(P_i)$ לא תספק את $G(P_i)$.

נשים לב שאם ל G יש את שלוש התכונות הנ"ל אז הפעלת G על בעיה נתונה מהסוג $q-csp$, מספקת את הרדוקציה מבעיית הפער $q-csp$ $\left(1,\frac{1}{3}\right)$ לבעיית הפער $(1,1-\varepsilon)\;E3-SAT$ עבור $\varepsilon=\frac{2}{3c}$.

לתרגם נוסחה ספיקה מ $q-csp$ לנוסחה ב $E3-SAT$ שהספיקות שלהם שקולה זה לא קשה. הבעיה היא להתמודד עם הנוסחאות שהן ספיקות באופן חלקי, או בפרט - פחות מ $\frac{1}{3}$ ספיקות. תהי $\widetilde{A}$ השמה לנוסחה $G(I)$ וניח ש I היא פחות מ $\frac{1}{3}$ ספיקה. נתבונן ב $\widetilde{A}$ מצומצמת למשתני I, נקרא לזה השמה A. ידוע לנו שזה מספק פחות מ $\frac{1}{3}$ מההסגרים, כלומר היא אינה מספקת לפחות $\frac{2}{3}$ מההסגרים, ולכן אם נסתכל על האופן שבו G תרגמה את ההסגרים הללו (כל אחד תורגם ללכל היותר c הסגרים) אזי כל אחד מהם (לפי דרישה 3) הוסיף לנו הסגר אחד שאינו מסתפק ב $\widetilde{A}$, כלומר יש לפחות $\frac{2}{3}\cdot\frac{1}{c}$ הסגרים שאינם מסתפקים בנוסחה המתורגמת, כלומר הנוסחה היא לכל היותר $1-\frac{2}{3c}$ ספיקה. ■

נותר להראות G שמבטיחה את שלושת התנאים.

לכל P_i נתחיל בלהציג את P_i ע"י נוסחת $q-CNF$, כלומר ע"י נוסחה מהסוג:

$$\left(\underbrace{... \vee ... \vee ... \vee ...}_{q-variables}\right) \wedge \left(\underbrace{... \vee ... \vee ... \vee ...}_{q-variables}\right) \wedge \left(\underbrace{... \vee ... \vee ... \vee ...}_{q-variables}\right) \wedge ...$$

איך עושים זאת? לכל השמה שאינה מספקת את P_i - נבנה הסגר שמכיל בדיוק את ההשמה הזו.

בשלב הבא נתרגם זאת ל $3-CNF$ בדרך אינטואיטיבית שנציג ע" דוגמה:

$$(x_1 \vee \overline{x_2} \vee x_3 \vee x_4) \rightarrow (x_1 \vee \overline{x_2} \vee x_n) \wedge (x_3 \vee x_4 \vee \overline{x_n})$$

ובתרגום הראשון שילמנו מחיר של ניפוח ב 2^q לכל היותר ובתרגום השני נשלם (בערך) q ולכן נקבל $c=O\left(q\cdot 2^q\right)$.

משפט 3.6 בעיית הפער IS $\left(\frac{1}{3},\frac{1}{3}-\varepsilon\right)$ (קבוצה בת"ל בגרף) היא בעיית NP-קשה עבור $\varepsilon>0$ ידוע כלשהו.

הוכחה: על ידי רדוקציה מ $3-SAT$, כפי שאנחנו כבר מכירים מחישוביות - ממירים כל הסגר בשלושה קודקודים בגרף הקשורים זה לזה (וכן כל משתנה ושלילתו) ונקבל שיש קבוצה בגודל $\frac{1}{3}$ מהגרף שהיא בלתי תלויה אם"ם הנוסחה ספיקה. ■

משפט 3.7 לכל $c > 0$ קבוע קיים $\beta(c)$ (קבוע שתלוי ב c) כך שבעיית הפער $(\beta, c \cdot \beta) - IS$ היא NP שלמה.

הגדרה 3.8 בהנתן שני גרפים לא מכוונים G, H נגדיר את המכפלה שלהם $G \times H$ ע"י:

$$V(G \times H) = \{(u, v) : u \in G, v \in H\}$$

$$((u_1, v_1), (u_2, v_2) \in E(G \times H)) \Leftrightarrow (u_1, u_2) \in E(G) \vee (v_1, v_2) \in E(H)$$

$$\alpha(G) = \frac{\max \ size\, of\, IS(G)}{|V(G)|}$$

למה 3.9

$$\alpha(G \times H) \geq \alpha(G) \cdot \alpha(H)$$

הוכחה: ניקח את S, T להיות קבוצות בת"ל מקסימליות ב G, H בהתאמה. נתבונן בקבוצת הקודקודים ב $G \times H$ הנתונה ע"י $S \times T$, ושם לכל (u_1, v_1) ו (u_2, v_2) אין צלעות ביניהן (מאי התלות של הקבוצות בגרפים המקוריים) ולכן זו קבוצה תלויה ב $G \times H$ שהיא ממש בגודל מכפלת הקבוצות התלויות, עם זאת - מכפלת הגרפים הטנזורית היא לכל היותר בגודל המכפלה שלהם כקבוצות, ולכן הלמה. ■

למה 3.10

$$\alpha(G \times H) \leq \alpha(G) \cdot \alpha(H)$$

הוכחה: תהי $U \subseteq V(G \times H)$ ונניח $|U| > \alpha(G) \cdot \alpha(H) \cdot |V(G \times H)|$ ויהיו S, T ההיטלים של U על קודקודי G, H בהתאמה (בנפרד). נובע או ש

$$|S| > \alpha(G) \cdot |V(G)|$$

או

$$|T| > \alpha(H) \cdot |V(H)|$$

ולכן S או T אינן בלתי תלויות, ומכאן U בלתי תלויה (כי במקום שבו ל S או ל T תהיה צלע - תיווצר צלע גם ב U) ולכן

$$\alpha(G \times H) \leq \alpha(G) \cdot \alpha(H)$$

■

מסקנה 3.11

$$\alpha(G \times H) = \alpha(G) \cdot \alpha(H)$$

מדוע זה מוכיח את המשפט? כי בהנתן גרף - ניצור את המכפלה הטנזורית שלו עם עצמו, ונשאל את אותה שאלה, וכאן החלק היחסי של הקבוצה הבלתי תלויה הוא ריבוע החלק שהיה, ולכן אם היתה קבוצה בת"ל קטנה מ $\frac{1}{3} - \varepsilon$, עכשיו היא קטנה מ $\left(\frac{1}{3} - \varepsilon\right)^2$, ובאופן דומה לקבוצה בת"ל גדולה מ $\frac{1}{3}$. נפעיל זאת k פעמים ונקבל בעיית פער $\left(\left(\frac{1}{3}\right)^k, \left(\frac{1}{3} - \varepsilon\right)^k\right)$. מכאן נובע שנוכל להקטין את הפער כרצוננו. נשים לב שההקטנה כאן היא רק ההקטנה של היחס בין החסמים, אבל למעשה שניהם קטנים מאד. מעניין גם לציין שאם מסתכלים על הבעיה המשלימה - היכולת שלנו לקרב דווקא נפגעת מהחזרה על הרדוקציה.

טענה 4.1 נזכיר את הבעיה מאחד השיעורים הקודמים: $PCP_{1,\frac{1}{3}}[r,O(1)] \subseteq P$ אם $r(n) = o(\log n)$.

הוכחה: חלקית - נראה ש $PCP_{1,\frac{1}{3}}[r,O(1)] \subseteq \bigcap_{c>0} NTIME(n^c)$.

אם $L \in PCP_{1,\frac{1}{3}}[r,O(1)]$ אז ניתן לייצר בזמן פולינומי את A, פרוטוקול PCP שנבצע כך: בהנתן n ביטים קלט, ו $r(n)$ ביטים מוגרלים, נשאל c (קבוע) שאלות את האורקל.

בהנתן I (קלט) ניצר לכל מחרוזת אקראית R את הפרדיקט הבוליאני P שמורכב מ c שאלות אל האורקל. נסמן זאת ע"י $P_R(q_1(R),...,q_c(R))$.

מקבלים בעיית $c-csp$, נסמן את קבוצת הפרדיקטים שעשויים להתקבל מהפרוטוקול A על הקלט I (לכל R מחרוזת ביטים אקראית אפשרית) ע"י $A(I)$.

ידוע כי אם $I \in L$ אז $A(I)$ ספיקה (במובן שכל הפרדיקטים ספיקים ע"י פונקציה מסויימת שמבצעת השמה למשתנים), ואם $I \notin L$ אז $opt(A(I)) \leq \frac{1}{3}$, כלומר ההשמה הכי מוצלחת לא תספק יותר משליש מהפרדיקטים. ואז מתקבל

$$|A(I)| = 2^{o(\log n)} \left[< n^{\delta} \forall constant\, \delta > 0\right]$$

נעבור על כל ההשמות האפשריות ל $A(I)$, כאלה יש $2^{2^{o(\log n)}}$, ונוכל להכריע לגבי השייכות של I ל L. אבל כמובן מכך ינבע

$$NP \subseteq PCP_{1,\frac{1}{3}}[o(\log n), O(1)] \subseteq TIME\left(2^{2^{o(\log n)}}\right) \subseteq \bigcap_{\delta>0} TIME\left(2^{n^{\delta}}\right)$$

אבל כיום אנחנו מניחים שקיים $\exists \delta_x$ כך ש:

$$NP \nsubseteq 2^{n^{\delta_x}}$$

■

אבל בעצם זה לא מוכיח את המבוקש, ננסה להוכיח את הטענה המקורית, והיא - שהדבר מוכל ממש ב P. **הוכחה:** קודם התחלנו מ $L \in NP$ כלשהי וקיבלנו רדוקציה ל $c-csp$, כאשר אם הקלט היה בגודל n - הפלט היה בגודל $2^{o(\log n)}$ שהוא תת-פולינומי לכל פולינום. זה נחמד מאד כי אפשר לעשות את הרדוקציה שוב (כי $c-csp$ עצמו ב NP)

$$I \underbrace{\rightarrow}_{reduction} A(I) \underbrace{\rightarrow}_{reduction} A_1(A(I))$$

כאשר $A_1(A(I))$ מתקבל מפרוטוקול PCP עם מספר תת-לוגריתמי של ביטים אקראיים על $c-csp$ (כי הבטחנו ש $PCP_{1,\frac{1}{3}}[o(\log n),1] \subseteq NP$), כשממשיכים מקבלים $A_1(A_1(A(I)))$ כי עכשיו אנחנו כבר מפעילים את אותו הפרוטוקול שוב ושוב (זה המתאים ל $c-csp$), ולכן לאחר k הפעלות מתקבל $A_1^k \circ A(I)$.

נעיר כי כבר ברדוקציה הראשונה - הפלט קטן לפחות בחצי - מדוע? כי מקבלים פלט בגודל $2^{o(\log n)}$, ועבור n גדול מספיק - זה קטן משורש n, שהוא כמעט תמיד קטן מחצי n, זה נכון לכל רדוקציה בשרשרת, ולכן אחרי מספר לוגריתמי של רדוקציות נקבל בעיית $c-csp$ קטנה מאד, כלומר פרדיקט יחיד שהוא ספיק אם"ם התחלנו עם בעיה $I \in L$.

נחשב את זמן הריצה של הרדוקציות: במקרה הגרוע ביותר ביצענו את הרדוקציה הראשונה ואחריה log פעמים את הזמן שנדרש לרדוקציה השניה. ולכן אם לרדוקציה הראשונה הזדקקנו לזמן של n^j, ולשניה הזדקקנו ל n^t כלשהו, אז בסך הכללי הזדקקנו ל:

$$n^j + n^t \log n$$

וזה פולינומי, מכאן $NP \subseteq P$ כנדרש. ■

נזכיר שפרוטוקול ה PCP הגדיר את פונקציית האורקל $O: Q \rightarrow \{0,1\}$ (Q היא קבוצת השאילתות לאורקל), אבל אפשר להגדיר $O: Q \rightarrow \{0,1\}^d$, ואז נסמן:

$$PCP_{c,s}[r,q,d]$$

עד כה הנחנו $d=1$, וגם בהמשך, אם לא נציין, הכוונה תהיה $d=1$.
נשאלת השאלה האם

$$NP \subseteq PCP_{1,1-\varepsilon}\left[O\left(\log n\right),1,O\left(1\right)\right]$$

?

נניח שכן, אם זה היה נכון - אז היינו מייצרים בעיית $1-csp$ בדומה לשאלה הקודמת ופותרים אותה, וזה כמובן ב P.

שעורי בית:

1. הוכח שלכל $\varepsilon>0$ מתקיים $PCP_{1,1-\varepsilon}\left[O\left(\log\left(n\right)\right),2,1\right]\subseteq P$:
באופן דומה לשאלה הקודמת, בהנתן בעיה I שיש להכריע האם היא מוכלת בשפה $L\in PCP_{1,1-\varepsilon}\left[O\left(\log n\right),2,1\right]$, נבנה בעיית $2-csp$ מכל תשובות האורקל האפשריות, נשים לב שבכל הסגר יש שני ליטרלים (שכל אחד יכול להיות אמת/שקר), ושמספר ההסגרים הוא 2^r כלומר $2^{O(\log n)}=n^{O(1)}$, קיבלנו בעיית $2-csp$ באורך שהוא פולינומי ב n. מספיק להציג פתרון פולינומי לבעיה זו ונקבל את המבוקש, בכל קירוב שנדרוש (כיוון שאפילו אם הסגר אחד יהיה שגוי - נתפוס את השגיאה ונשיב בשלילה על השאלה "$I\in L$?").
2. הוכח שקיים $\varepsilon>0$ כך ש $NP\subseteq PCP_{1,1-\varepsilon}\left[O\left(\log\left(n\right)\right),2,3\right]$

קבוצה בלתי תלויה ($I.S.$)

תהי $L\in PCP_{c,s}\left[r,q\right]$, ויהי I קלט שעלינו להכריע האם $I\in L$ או $I\notin L$.
יהי P פרוטוקול כנ"ל עבור L.

הגדרה 4.2 P מייצר $Q_1,...,Q_q$ שאילתות, מתקבל מהאורקל רצף של q ביטים $O\left(Q_1\right),O\left(Q_2\right),...,O\left(Q_q\right)$ והפרוטוקול מייצר פרדיקט בוליאני שנסמנו:

$$P_{R,I}\left(O\left(Q_1\right),O\left(Q_2\right),...,O\left(Q_q\right)\right)$$

נאמר שהתהליך שתיארנו הוא רשומון($transcript$) של הפרוטוקול, ומספר הרשומונים האפשריים הוא לכל היותר $2^r\cdot 2^q$. נאמר שרשומון מקבל אם הפרדיקט נותן ערך אמת.

.

הגדרה 4.3 נאמר ששני רשומונים הם עקביים אם לכל שאילתה שמופיעה בשניהם האורקל משיב באותו האופן.
(בעצם כל שני רשומונים של אותו אורקל - הם עקביים, נעשה שימוש בעובדה זו בהמשך)

בהנתן הקלט I, נוכל לייצר גרף $G\left(I\right)$, שקודקודיו הם קבוצת כל הרשומונים המקבלים וזוג רשומונים $\left(t_1,t_2\right)$ מהווה צלע אם"ם הזוג הוא בלתי עקבי. למעשה נייצר גרף של כל הרשומונים בכלל, אבל נחבר את כל הרשומונים שאינם מקבלים זה לזה (נשתמש בזה בהמשך כיוון שהם לא יוכלו להיות חלק משום קבוצה בת"ל) וכך נקבל $\left|G\left(I\right)\right|=2^{r+q}$.
נניח ש $I\in L$: לפי PCP במקרה כזה קיים אורקל O שמספק את החלק הc מתוך הפרדיקטים שמקורם בהגרלת r ביטים, ולכן קיימת ב G_I קבוצה בלתי תלויה בגדול $c\cdot 2^r$.
נניח ש $I\notin L$: נוכיח שכל קבוצה בלתי תלויה $S\subseteq V\left(G_I\right)$ היא עקבית ולכן ניתן לבנות ממש בידיים אורקל O שמתאים לה, ושאלו יהיו התשובות שיתן. אבל על פי הנחת PCP מתקיים שלא ניתן לייצר אורקל שמשיב נכונה על יותר מהחלק ה s מתוך הפרדיקטים, ולכן גודל הקבוצה הבלתי תלויה הוא די קטן: $\left|S\right|\leq s\cdot 2^r$.
האם קיבלנו רדוקציה מבעיית הקבוצה הבת"ל לבעיית פער?
כן, קיבלנו בעיית פער $\left[\frac{c}{2^r},\frac{s}{2^q}\right]$ (זה מצטמצם ל $\left[\frac{c\cdot 2^r}{2^{r+q}},\frac{s\cdot 2^r}{2^{r+q}}\right]$), זמן הרדוקציה יהיה $poly\left(n,2^{r+q}\right)$ ועבור q,r מספיק נוחים - זה יהיה פולינומיאלי ואז מ $NP\subseteq PCP_{1,1-\varepsilon}\left[\log n,1\right]$ נקבל רדוקציה מכל $L\in NP$ לבעיית הפער $gIS-\left[\frac{1}{2^q},\frac{\varepsilon}{2^q}\right]$.
הערה: אם קיים קירוב $1-\varepsilon$ בזמן פולינומיאלי לקבוצה בת"ל, אז:

$$NP\subseteq PCP_{1,1-\varepsilon}\left[\log n,1\right]\subseteq P$$

נוכל לחזור על הרדוקציה כדי לקרב את IS שוב ושוב, נעשה זאת $\log\log n$ פעמים. למה? בפעם הראשונה נסתמך על ההכלה הידועה לנו ש $PCP_{1,\varepsilon}[r,q] \subseteq PCP_{1,\varepsilon^2}[2r,2q]$ ובהמשך באופן רקורסיבי אחרי k הפעלות נקבל $PCP_{1,\varepsilon^{2^k}}\left[2^k r, 2^k q\right]$ ואם התחלנו עם $r=\log n$ ו $q=O(1)$ אז נקבל אחרי $\log\log n$ הפעלות

$$PCP_{1,\varepsilon^{\log n}}\left[\log^2 n, \log n\right]$$

כאשר זמן הריצה יהיה $n^{\log n}$ (ניתן לחשב זאת לפי ההכלה האחרונה, שזמן הריצה שלה הוא פולינומיאלי ב $\left(n, 2^{\log^2 n + \log n}\right)$, ולכן אם נמצא אלגוריתם פולינומיאלי שמקרב את IS עד כדי $\varepsilon^{\log n}$, אז $NP \subseteq TIME\left(N^{\log n}\right)$, כי הרי התחלנו מבעיית NP ועשינו רדוקציה לבעיית קירוב של IS.

5 PCP

כזכור ראינו משפט שאם קיים אלגוריתם בזמן פולינומיאלי שמקרב את בעיית IS עד כדי גורם $\frac{c}{s}$, אז

$$PCP_{c,s}[r,q] \subseteq TIME\left(Poly\left(n, 2^{r+q}\right)\right)$$

נזכיר שמלמת החיזוק קיבלנו

$$PCP_{c,s}[r,q] \subseteq PCP_{c^2,s^2}[2r,2q]$$

ולכן אפשר להקטין את הגודל $1-\varepsilon$ כרצוננו ולקבל:

$$\forall \underbrace{\delta}_{constant} \quad NP \subseteq PCP_{1,\delta}\left[\log(n), 1\right]$$

מסקנה 5.1 אלא אם $NP \subseteq P$, לכל $\delta > 0$ לא קיים אלגוריתם שרץ בזמן פולינומיאלי ומקרב את IS עד כדי δ.

הערה 5.2 נזכיר את הבעיה הדואלית ל IS והיא כיסוי קודקודים VC, נשים לב שגם אם נקבל קבועים α, β שמקרבים את בעיית ה IS - וגם אם היחס ביניהם יהיה גדול, היחס בבעיה הדואלית יהיה $\frac{1-\alpha}{1-\beta}$, שיכול להיות מאד קרוב ל 1, ולכן זה לא נותן לנו התקדמות. מה שכן, הוכח שקירוב של יותר מ ...1.36 לבעיית כיסוי הקודקודים הוא $NP-Hard$ (ספרא ודינור), כמו כן עודד רגב ואחרים הראו שבהנחת UGC $(Unique\,Games\,Conjecture)$ לקרב את VC ביותר מ $2-\delta$ זה $NP-Hard$.

שעורי בית

1. הראו שקיים ε כך שקירוב $1+\varepsilon$ של בעיית כיסוי הקודקודים היא בעיה $NP-Hard$. (בהנחת UGC).

עבור $\alpha = \alpha(n)$ שהיא פונקציה לכל היותר פולינומית:

משפט 5.3

$$PCP_{c,s}[r,q] \subseteq PCP_{c^\alpha, S^\alpha}[\alpha r, \alpha q]$$

פשוט חוזרים על הפרוטוקול $\alpha(n)$ פעמים.

מסקנה 5.4 נבחר $\alpha = \log^\beta n$ ונקבל:

$$NP \subseteq PCP_{1,1-\varepsilon}[\log n, 1] \subseteq PCP_{1,(1-\varepsilon)^\alpha}\left[\log^{(1+\beta)} n,\ \log^\beta n\right]$$

ולכן לא קיים קירוב שרץ בזמן פולינומיאלי לבעיית IS עד כדי $(1-\varepsilon)^{\alpha} \approx (const.)^{\log^{\beta} n}$ בגרפים שגודלם $n^{O\left(\log^{\beta} n\right)}$.
אלא אם $NP \subseteq TIME\left(2^{\log^{(1+\beta)} n}\right) = TIME\left(n^{O\left(\log^{\beta} n\right)}\right)$.
נשים לב:

$$2^{\log^{\beta} n} = \left(2^{\log^{1+\beta} n}\right)^{\frac{1}{\log n}} = \left(2^{\log^{1+\beta} n}\right)\left(\log\left(2^{\log^{1+\beta} n}\right)\right)^{\frac{1}{1+\beta}}$$

אם גודל הגרף הוא $N = 2^{\log^{1+\beta} n}$ אז נקבל שזה $NP - Hard$ לקרב את IS עד כדי פקטור של $\sqrt[\frac{1}{(\log N)^{\frac{1}{1+\beta}}}]{N}$.

למת חיזוק #2

למה 5.5 $Amplification$:

$$NP \subseteq PCP_{1,(1-\varepsilon)^k}\left[\log n + k,\, k\right]$$

מסקנה 5.6 IS קשה לקירוב בטווח של n^{β} עבור β קבוע כלשהו, אלא אם $NP \subseteq P$.

גרפים מרחיבים

הגדרה 5.7 גרף G, d־רגולרי, הוא גרף מרחיב אם:

$$\forall S \subseteq V(G)\,,\ |S| \leq \frac{1}{2}|V(G)| \rightarrow \underbrace{|E(S, V(G) \setminus S)|}_{edges\ going\ out\ of\ S} \geq |S|$$

מסתבר שלא פשוט לבנות גרפים כאלה עבור d קבוע בגודל גדול כרצוננו, ולמעשה זה בלתי אפשרי כשהגרף הוא גם מישורי.

משפט 5.8 קיימת משפחה מפורשת חזק של גרפים מרחיבים d־רגולריים.

הגדרה 5.9 משפחה של גרפים $G_1, ..., G_k, ...$ היא מפורשת $(explicit)$ אם קיים אלגוריתם A כך שלכל n מייצר $A(n) = G_n$, $|V(G_n)| = n$ כך ש A רץ בזמן פולינומי.
משפחה של גרפים תיקרא מפורשת חזק $(strongly\, exlpicit)$אם קיים A^S אלגוריתם, שלכל n, i, j מחזיר את $A^S(n,i,j)$=האינדקס של השכן ה j של הקודקוד ה i בגרף G_n, ו A^S רץ בזמן $Poly(\log n)$. כמובן אם משפחה היא מפורשת חזק ־ בפרט היא מפורשת, שכן ניתן לבנות לכל k את הגרף G_k פשוט ע"י מעבר סדרתי על כל ה i וה j.

משפט 5.10 יהי G גרף מרחיב d רגולרי. תהי $B \subseteq V(G)$ ונסמן $\beta = \frac{|B|}{|V(G)|}$, אז קיים קבוע α (שהוא פונקציה של הדרגה d), כך שלכל k, ההסתברות שהילוך מקרי בן k צעדים על G מוכל ב B, קטנה מ $\beta^{\alpha \cdot k}$. נגדיר הילוך מקרי ע"י בחירת קודקוד ב G בהתפלגות אחידה, ונצעד מקודקוד לאחד משכניו בהתפלגות אחידה גם כן. נאמר שהילוך מקרי מוכל ב B אם כל אחד מהקודקודים בהילוך שייך ל B.

משמעות הדבר שבאופן עקרוני הילוך מקרי, אפילו אם התחיל ב B במקרה ־ בסיכוי טוב יצא מ B. למעשה ־ ההסתברות להשאר ב B בכל צעד, קרובה להסתברות להתחיל ב B מלכתחילה.
נחזור להוכחת למת החיזוק השניה:

$$NP \subseteq PCP_{1,(1-\varepsilon)^k}\left[\log n + k, k\right]$$

בהנתן I שצריך להכריע אם הוא שייך ל $L \in NP$, נפעיל את הפרוטוקול $PCP_{1,1-\varepsilon}\left[\log n, 1\right]$ המקורי שאנחנו יודעים שקיים. נייצר גרף מרחיב G_n כלשהו, המובטח מהמשפחה שהיא מפורשת חזק. נגריל קודקוד v בתוך G_n (באמצעות ה $\log n$ ביטים המובטחים לנו), ונגריל כמה ביטים נוספים (מספר חסום, תלוי ב d) שיכריע לאיזה קודקוד משכניו של

v נטייל. נניח שטיילנו ל u - אזי הוא מסומן על ידי $\log n$ ביטים כלשהם, נשתמש בהם בתור קלט לאלגוריתם המקורי - ונפעיל אותו שוב. נמשיך כך לטייל בגרף, וכמובן שנקבל רק אם במשך כל הטיול קיבלנו תשובות מספקות מהאורקל. נגדיר את הקבוצה הרעה (שאנחנו מקווים לא להיות בתוכה) הוא:

$$B(I,O) = \{r : P(r,I,O) = acc\}$$

(כאשר אנחנו מייצגים כל קודקוד על ידי $\log n$ביטים גם כאן). כפי שציינו במשפט - ההסתברות להישאר כל הזמן בתוך הקבוצה B (שהיא קבוצה "רעה" שעבורה נקבל acc כוזבים) תהיה קטנה מאד.

PCP 6

ראינו בשבוע שעבר חיזוק למשפט ה PCP בעזרת הילוך מקרי על גרף מרחיב:

$$NP \subseteq PCP_{1,\left(\frac{1}{2}\right)^k}\left[O(\log n + k), O(k), 1\right]$$

עבור פרוטוקול PCP שמבצע q שאילתות, ומקבל תשובה באורך d ביטים לכל שאילתה, עם נאותות s, נגדיר את סיבוכיות השאילתה ע"י

$$\frac{q \cdot d}{\log\left(\frac{1}{s}\right)}$$

מערכת ה PCP שהצגנו עבור NP היא בעלת סיבוכיות שאילתה קבועה, לכל מספר שאילתות שנרצה (כל עוד הוא פולינומי).

בנוסף, אנחנו יכולים לקבל מספר שאילתות גדול מ $\Omega(\log n)$, אם נדאג שמספר הביטים האקראיים הוא $O(\log n)$.

נניח בשלילה שאנחנו יכולים לשפר את הנאותות מבלי לפגוע במספר השאילתות וההגרלות, במילים אחרות - שניתן לקבל:

$$NP \subseteq PCP_{1,\left(\frac{1}{2}\right)^{f(k)}}\left[O(\log n + k), O(k), 1\right]$$

כאשר $f(k)$ היא פונקציה שהיא ממש יותר מליניארית. אם כך היה המצב, אז היינו בוחרים את k להיות $k = \Theta(\log n)$, אז נקבל:

$$s < \frac{1}{n^{\omega(1)}}$$

כי הנאותות לפי הנחתנו קטנה ממש יותר מהר מ $\frac{1}{2}^k$.
מתקבל.

$$q \sim \log n$$

נזכיר שהיתה לנו הרדוקציה באמצעות הגרף המרחיב

$$PCP_{1,\varepsilon} \longrightarrow gap[1,\varepsilon]\ q - csp$$

מכיוון ש $q \sim \log n$ אז $\varepsilon = \frac{1}{n^{\omega(1)}}$, ומכיוון שהרדוקציה לקחה לנו זמן פולינומיאלי ב $n, 2^r$ וכעת $r = O(\log n)$, כלומר הרדוקציה לוקחת זמן פולינומיאלי ב n.

מהרדוקציה נקבל מספר פולינומיאלי של פרדיקטים כך שאם הבעיה היתה ב PCP הנתון, אז כל הפרדיקטים ספיקים, ואחרת לכל היותר החלק ה $\frac{1}{n^{\omega(1)}}$ מהם ספיקים. אבל כאמור מספר הפרדיקטים הוא פולינומיאלי ב n, ואז החלק ה $\frac{1}{n^{\omega(1)}}$ מהם הוא פחות מאחד. מכאן נובע שאו שכל הפרדיקטים ספיקים או שאף אחד מהם אינו ספיק. במקרה כזה מספיק לבחור פרדיקט ולבדוק אותו, ו $L \in PCP$ אם"ם הפרדיקט ספיק. זו נוסחה מעל $\log n$ משתנים, ולכן בדקת כל ההצבות לה היא פולינומיאלית ב n, ומכאן נובע שאנחנו יכולים להכריע בזמן פולינומי האם $L \in NP$ או לא. לכן אם ניתן לקבל נאותות כנ"ל אז $NP \subseteq P$.

עד כה כאשר שיפרנו את הנאותות, תמיד שילמנו במספר השאילתות (שהיה קבוע כפול המעריך בנאותות). נרצה להחזיק את מספר השאילתות קבוע ועם זאת לשפר את הנאותות.

משפט 6.1 $[DFKRS]$

$$\forall \beta < 1 \, \exists q = q(\beta) \; s.t. \; NP \subseteq PCP_{1,\varepsilon}\left[O(\log n), q, d\right]$$

כך ש

$$d = \log^{\beta} n, \; \varepsilon = 2^{-d}$$

כלומר בהנתן β, ניתן על ידי מספר שאילתות קבועות ומספר ביטים לשאילתה שהוא קצת יותר קטן מלוגריתמי ב n, לקבל נאותות שהיא כמעט פולינומיאלית ב n ($2^{\log^{\beta} n}$) .

6.1 PCP עם שתי שאילתות בלבד

הגדרה 6.2 משחק סיבוב 1, עם 2 מוכיחים. $(2\,prover\,1\,round\,game)$ $[BGKW1988]$
נניח שאליס ובוב הם שני המוכיחים, המשחקים נגד גורם שלישי (הטבע, אלוהים, וכו'). אליס ובוב ממוקמים בשני חדרים נפרדים כך שאין להם יכולת לתאם עמדות.
חוקי המשחק מוגדרים על ידי התפלגות על זוגות של שאלות D ופרדיקט P (די גדול) על שאלות המשחק, שניהם ידועים לאליס ובוב.

הפרדיקט P מוגדר על זוג של שאלות ושתי תשובות כך ש $P\left(\underbrace{x,y}_{questions} \;,\; \underbrace{a,b}_{answers}\right)$ נותן "זכיה" או "הפסד" לכל צירוף כזה.
מהלך המשחק הוא:

1. מגרילים מההתפלגות D זוג שאלות (x,y)
2. שולחים את x לאליס ואת y לבוב.
3. אליס משיבה a ובוב משיב b.
4. לפי P מתקבלת החלטה האם אליס ובוב זכו או הפסידו.

בהנתן אסטרטגיה עבור אליס ובוב - יש להם הסתברות לזכיה, נסמן ע"י $v(G)$ את ההסתברות המקסימלית לזכיה (ערך המשחק בעצם).

6.1.1 דוגמה:

x, y הם ביטים אקראיים (בהסתברות חצי וכו') בלתי תלויים.
התשובות הן איברים בקבוצה $\{A,B\} \times \{0,1\}$ (כלומר שם של אחד המשתתפים וביט).
פרדיקט הזכיה: אליס ובוב זוכים אם ורק אם בתשובה של שניהם יש אותו השם X ואותו הביט j, וגם הביט j הוא ביט השאלה שנשלח למשתתף X.
אסטרטגיה אפשרית: שני הצדדים שולחים B, בוב שולח את הביט שהוא קיבל ואליס שולחת 1. אם באמת בוב קיבל את הביט 1 - הם ניצחו, ואם לא - הפסידו. מכאן שהסתברות הזכיה באסטרטגיה זו היא $\frac{1}{2}$. באותו אופן ניתן לבחור מראש שתמיד שניהם ישיבו $B,1$ וגם כך נקבל הסתברות זכיה של $\frac{1}{2}$. לא קשה להשתכנע שאי אפשר לשפר זאת.

הגדרה 6.3 לשפה L יש פרוטוקול $MIP_{c,s}[2,1]$ אם בהנתן I אפשר לייצר בזמן פולינומיאלי משחק של 2 מוכיחים וסיבוב 1, שהוא $G(I) = G$ כך שמתקיים:

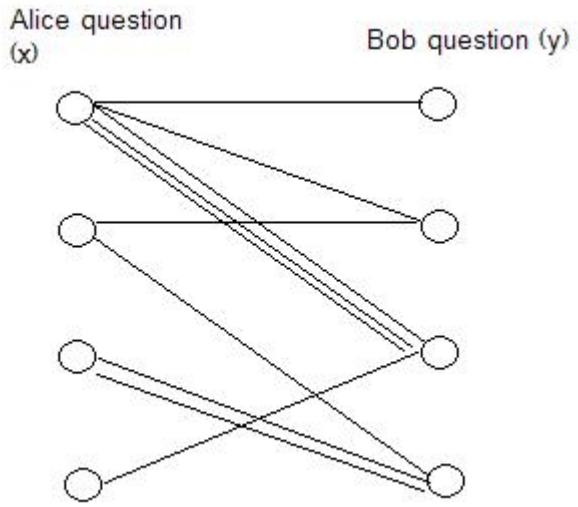
1. אם $I \in L$ אז $v(G) \geq c$
2. אם $I \notin L$ אז $v(G) \leq s$

6.1.2 מה הקשר ל PCP?

אליס ובוב משמשים בתור האורקל, כאשר אנחנו משחקים נגדם. אם $I \in L$ אז אליס ובוב ינצחו בהסתברות גדולה מ c, ואם $I \notin L$ אז הם ינצחו בהסתברות קטנה מ c.
לכן נשים לב:

$$MIP_{c,s}[2,1] \subseteq PCP_{c,s}[O(\log n), 2, ?]$$

מדוע? מה יעשה פרוטוקול ה PCP שלנו? בהנתן I, נייצר את המשחק $G(I)$, נבחר זוג שאלות $(x,y) \sim D(G(I))$, והשאלה הראשונה תהיה $A(x)$=מה תשובתה של אליס לשאלה x?
השאלה השניה תהיה $B(y)$=מה תשובתו של בוב לשאלה y.
לכאורה, מכיוון ששאלנו על בוב אחרי ששאלנו על אליס, לא נשמרה ההפרדה בין המידע שבידי אליס למידע שבידי בוב, אבל האורקל הוא בסך הכל פונקציה, ולכן הערכים נקבעים מראש ולא תלויים בשאלות קודמות.
קל לראות שהשלמות c והנאותות s.
אם נגדיר שתיאור ההתפלגות הוא ממש רשימת כל הזוגות (כאשר כולם בהסתברות שווה, ואם רוצים לאחד מהם הסתברות גבוהה יותר, רושמים אותו מספר פעמים) אז מכיוון שהבטחנו שייצור המשחק הוא פולינומיאלי בגודל של $n = I$, נקבל שמספר הזוגות הוא פולינומיאלי ב n גם כן, ולכן כדי להגריל זוג צריך $O(\log n)$ ביטים, ולכן זהו מספר הביטים האקראיים בפרוטוקול לעיל.
אפשר לתאר את המשחק והאסטרטגיה המיטבית ע"י גרף דו צדדי, כאשר בצד אחד השאלות לבוב (ותשובותיו) ובצד השני השאלות לאליס (ותשובותיה), ההסתברות לכל צלע היא אחידה, כמקודם, אם נרצה התפלגות אחרת - אז נוכל להוסיף צלעות בין זוגות של שאלות.



כל צלע למעשה מייצגת רביעיה x, y, a, b ולכן לפי תוצאת הפרדיקט P - אפשר גם לסמן כל צלע ע"י "האם היא מסמלת זכיה או הפסד".
אם

$$NP \subseteq MIP_{c,s}[2,1]$$

כאשר גודל התשובות הוא d, אז בעיית הפער $[c,s]$ עבור $2 - csp(d)$ היא NP-שלמה, הסימון משמעו שזו נוסחא שבכל פרדיקט יש 2 משתנים, ואורך כל השמה לאחד המשתנים היא d ביטים.

משפט 6.4 קיים $\varepsilon > 0$ כך ש $NP \subseteq MIP_{1,\varepsilon}[2,1]$ כאשר התשובות הן בגודל קבוע.

הוכחה: נעבוד על ידי רדוקציה מבעיית הפער $[1, 1-\varepsilon']$ עבור $3-SAT$. אזי או שכל ההסגרים ספיקים, או שלכל היותר $1-\varepsilon'$ מהם ספיקים. נרצה פרוטוקול MIP עבור שפה זאת. נציג את המשחק ובכך תתקבל הרדוקציה ל $MIP_{1,\varepsilon}[2,1]$ (לא אותו אפסילון).
המשחק:

1. בחר הסגר באקראי, נניח i, ואז

$$C_i = (X_{i_1} \vee X_{i_2} \vee X_{i_3})$$

2. שלח את C_i לאליס
3. בחר $j \in \{1,2,3\}$ באקראי
4. שלח את X_{i_j} לבוב.

נאמר שאליס ובוב זכו אם אליס מחזירה השמה מספקת ל C_i ובוב מחזיר ערך עבור X_{i_j} שהוא עקבי עם תשובתה של אליס.

נסמן את

אם $I \in 3-SAT$ אז $v(G(I)) = 1$.

אם I היא פחות מ $1-\varepsilon'$ ספיקה, אז $v(I) \leq 1-\varepsilon'$.

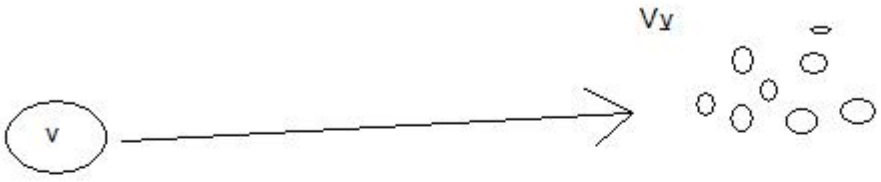
האסטרטגיה של בוב מגדירה בעצם השמה $\mathcal{A}$ עבור I, כך ש $v_{\mathcal{A}}(I) \leq 1-\varepsilon'$, ואז בהסתברות גדולה מ ε', או שתשובתה של אליס איננה מספקת או שאיננה עקבית עם $\mathcal{A}$. כלומר בהסתברות לפחות $\frac{\varepsilon'}{3}$ אליס ובוב מפסידים, ואז $\varepsilon = \frac{\varepsilon'}{3}$. ■

7 :

הגדרה 7.1 בעיית $Label-Cover$ - בהנתן בעיית $2-csp(7)$ (כלומר 2 משתנים בכל אילוץ שיכולים לקבל 7 ערכים כ"א), נאמר שהבעיה היא בעיית $Label-Cover$ אם בהנתן השמה למשתנה הראשון בכל אילוץ - קיימת השמה יחידה למשתנה השני אשר מספקת את האילוץ.

נרצה לבנות רדוקציה מ LC שהוזכרה לבעיית $E3-LIN(2)$ (בעיית תכנון ליניארי בשלושה משתנים, בשדה עם מציין 2).

נתרגם את המשתנה הראשון v באילוץ לקבוצה של משתניםV_v, ואת המשתנה השני u לקבוצה של משתנים V_u:



נרצה למת חיזוק לבעיית הפער עבור LC, מהרדוקציה נראה שמתקיים שבעיית הפער $\left(1-\delta,\frac{1}{2}+\varepsilon\right)$ עבור $E3-LIN\left(2\right)$ היא NP שלמה. לאחר מכן נראה רדוקציה מבעיה זו לבעיית הפער $\left(1-\delta,\frac{7}{8}+\varepsilon\right)$ עבור $E3-SAT$. בהמשך נשלים את התשתית התיאורטית הקושרה להוכחת ה PCP.

למת חיזוק ל MIP

נוכיח למת חיזוק עבור $MIP\left[2,1\right]$, נרצה שאם $v\left(G\right)=1$ אז אחרי החיזוק ישאר 1, אבל אם $v\left(G\right)=1-\varepsilon$ אז ערך המשחק יקטן.

נוכל לשחק את המשחק כמה פעמים בזו אחר זו ולהכריז על ניצחון רק אם התקיים ניצחון בכל המשחקים. הבעיה כאן היא שאז זה כבר לא עומד בהגדרה של $MIP\left[2,1\right]$, כי כביכול משחקים יותר מסיבוב אחד.

הגדרה 7.2 חזרה מקבילית $G\to G^{\otimes k}$ נבחר k זוגות $\left(x_1,y_1\right),\left(x_2,y_2\right),...,\left(x_k,y_k\right)$ באופן בלתי תלוי זה מזה (כלומר כל שני זוגות הם בת"ל זה בזה), ואז נשלח את השאלות $\overline{x}=\left(x_1,x_2,...,x_k\right)$ לאליס ו $\overline{y}=\left(y_1,y_2,...,y_k\right)$ לבוב. התשובות תהיינה $\overline{a}=\left(a_1,a_2,...,a_k\right)$ ו $\overline{b}=\left(b_1,b_2,...,b_k\right)$ בהתאמה.

נגדיר את הפרדיקט על התשובות להיות:

$$P^{\otimes k}\left(\overline{x},\overline{y},\overline{a},\overline{b}\right)=\prod_{i=1}^{k}P\left(x_i,y_i,a_i,b_i\right)$$

בעצם כאילו שיחקנו את כל המהלכים של השחקן הראשון במכה, ואז את כל המהלכים של השחקן השני ואז הכרענו.

מאמר של פורטנאו, סיפסר ועוד מישהו מ 1988קבעו כי "ברור ש"

$$v\left(G^{\otimes k}\right)=v\left(G\right)^{k}$$

וזו בדיוק התכונה שרצינו לגבי ערך המשחק.

אבל באותה שנה פורטנאו פירסם יחד עם ארטה מאמר שמראה שזה אינו נכון בעצם.

דוגמה נגדית:

נזכיר את המשחק מההרצאה הקודמת: x,y ביטים אקראיים, אליס ובוב מחזירים שם וביט, כך שהם זוכים רק אם הם מסכימים על השם והביט המוחזר וגם הביט שהחזירו הוא אכן הביט שנשלח למשתתף שאת שמו החזירו.

אנחנו יודעים ש $v\left(G\right)=\frac{1}{2}$, האם מתקיים $v\left(G^{\otimes2}\right)\leq\frac{1}{4}$?

אליס מקבלת x_1,x_2 ובוב מקבל y_1,y_2.

האסטרטגיה של אליס תהיה: התשובה הראשונה (אליס, x_1), התשובה השניה (בוב , x_1).

האסטרטגיה של בוב תהיה: התשובה הראשונה (אליס, y_2), התשובה השניה (בוב, y_2)

זוכים בכל המשחקים רק אם $x_1=y_2$, המשתנים הללו בלתי תלויים ולכן ההסתברות לכך היא $\frac{1}{2}$ וזה ערך המשחק.

הערה 7.3 הצגנו משחק שעבורו $v\left(G^{\otimes2}\right)=v\left(G\right)$, יש שאלה פתוחה אם קיים משחק שעבורו $v\left(G^{\otimes3}\right)=v\left(G\right)$.

ורביצקי הראה ב 94' כי אכן כאשר $k\to\infty$ מתקבל $v\left(G^{\otimes k}\right)\to0$, אבל האמצעים היו לא קונסטרוקטיביים, ולא היה ברור מכך איך בדיוק ערך המשחק קטן, ולכן זה לא שימושי לנו.

ב 95 הראה ... כי אם אליס ובוב עונים מתוך קבוצה בגודל שקטן מ l, ו $v\left(G\right)=1-\varepsilon$, אז מתקיים:

$$v\left(G^{\otimes k}\right)\leq\left(1-\varepsilon^{32}\right)^{\frac{C\cdot k}{\log\left(l\right)}}$$

לאחרונה שופר החסם ל $\left(1-\varepsilon^{3}\right)^{\frac{C\cdot k}{\log\left(l\right)}}$

בכל מקרה - למרות שכנראה צריך לנפח את התשובות מאד, מובטחת לנו ירידה מעריכית של ערך המשפט בשלב מסויים.

בבעיית ה $Lable-Cover$ שראינו, גודל קבוצת התשובות האפשרית הוא 7, ולכן מובטח לנו שעבור חזרה מקבילית k פעמים - יהיה $\left(1-\varepsilon^{3}\right)^{C\cdot k}$ (כאן הנחנו שהחלוקה בלוג 7 נבלעה ב C). ולכן בהנתן ε' קבוע נקבל:

$$\left(1-\varepsilon^{3}\right)^{C\cdot k}=\varepsilon'$$

$$C \cdot k \cdot \log\left(1-\varepsilon^3\right) = \log \varepsilon'$$

$$k = \frac{\log\left(\varepsilon'\right)}{C \cdot \log\left(1-\varepsilon^3\right)}$$

כעת קיבלנו בעיה שבה התשובות הן וקטורים באורך k שלכל קואורדינטה בהם יש 7 אפשרויות, ולכן זו בעיית $Label-Cover\left(7^k\right)$.

הרדוקציות הן פולינומיאליות. שהרי אם הייצוג של המשחק הוא על ידי גרף דו"צ כפי שראינו בשבוע שעבר, לא קשה להיווכח שאם חוזרים על תהליך ההגרלה והבדיקה k פעמים וכו' - מקבלים זמן ריצה n^k בערך, וכיוון ש k נקבע מראש לפי איכות הקירוב הדרוש - זה פולינומי ב n.

כדי להשלים את המשפט שאומר שקיים ε כך שבעיית הפער $(1,\varepsilon)$ עבור $LC\left(d\right)$ היא NP-שלמה, חסר להראות שכאשר חוזרים על משחק שיש לו את תכונת ההטלה (כלומר תשובתה של אליס קובעת באופן יחיד את האפשרות לתשובה נכונה של בוב), מקבלים משחק שיש לו את תכונת ההטלה, אך לא קשה להשתכנע בכך אם בוהים בחומר מספיק.

הגדרה 7.4 משחק יקרא ייחודי אם לכל שאלה x,y, לכל a קיים b יחיד (וגם לכל b קיים a יחיד) כך ש $P\left(x,y,a,b\right)$ נותן ערך אמת (=זכיה).

UGC 2002: לכל ε ו $\delta>0$, קיים $d\left(\varepsilon,\delta\right)$ כך שבעיית הפער $(1-\delta,\varepsilon)$ עבור $Label-Cover$ ייחודי, היא NP שלמה.

שעורי בית

1. הוכח שלהכריע האם משחק ייחודי הוא ספיק היא בעיה ב P
2. הוכח שאם $P\neq NP$, אז בלתי אפשרי להשיג פחות מסיבוכיות שאילתה קבועה (כלומר לקבל איזושהי פונקציה יורדת), כלומר שקיים קבוע כלשהו שאי אפשר להשיג פחות ממנו.

רדוקציה מבעיית פער LC ל $E3-LIN\left(2\right)$

נתבונן על הבעיה ב $(1,\varepsilon)-LC\left(d\right)$ (כלומר בעיית פער $Label-Cover$ עם d ערכים אפשריים), ונרצה לתרגם את המשתנים למשתנים בוליאניים, כמובן שאם נרצה לייצג את d ערכי המשתנה, נצטרך לפחות $\log d$ משתנים בוליאניים, למעשה נשתמש ב 2^d, שזה ניפוח מעריכי שנראה על פניו לא הכרחי, אבל אנחנו לא מכירים דרך לשפרו.

ואז את המשתנה v נתרגם לקבוצת משתנים בגודל 2^d, כל אחד מהם מתאים לאיבר בקבוצה $\{1,-1\}^d$. נייצג השמה למשתנים החדשים ע"י:

$$f_v : \{1,-1\}^d \to \{0,1\}$$

כך לכל משתנה x ב 2^d הביטים שמייצגים את v - ההשמה $f_v\left(x\right)$ תיתן לו את הערך המתאים כך שייצג את השמת הערך הנתון ל v . כלומר, אם למשל המשתנה v יכול לקבל 3 ערכים, אזי קיבלנו קבוצה בגדול 8 של משתנים בוליאניים, שהם:

$$000, 001, 010, 011, 100, 101, 110, 111$$

ואז, למשל, עבור $v=0$ נקבל שההשמה לשמונת המשתנים הללו היא $(0,0,0,0,1,1,1,1)$ עבור $v=1$ נקבל $(0,0,1,1,0,0,1,1)$ ועבור $v=2$ נקבל $(0,1,0,1,0,1,0,1)$. במילים אחרות - לוקחים את אחת הקואורדינטות בייצוג הבינארי, בפרט - את הקואורדינטה שמתאימה לאינדקס של הערך מתוך d הערכים שנתנו ל v.

8 :

נזכיר שאנחנו במהלך נסיון לרדוקציה מבעיית הפער $LC(k)$ עבור $(1,\varepsilon)$, לבעיית $E3-LIN(2)$. מכיוון שלהכריע האם בעיית $E3-LIN(2)$ ספיקה לחלוטין זה פולינומיאלי, ולכן נרצה שבעיית הפער תמיר בעיות ספיקות לחלוטין ב $LC(k)$ לבעיות שהן לפחות $1-\delta$ ספיקות ב $E3-LIN$, ובעיות שהן פחות מ ε ספיקות ב $LC(k)$ יומרו לבעיות שהן לכל היותר $\frac{1}{2}+\varepsilon'$ ספיקות ב $E3-LIN(2)$.

מה שיקרה יהיה שלפי בחירת ε', δ נבחר את ε, שבתורו יקבע את k.

נראה את הרדוקציה באופן מקומי:

בעיית ה LC מיוצגת ע"י גרף מכוון, כאשר כל צלע מייצגת פונקציה כלשהי מ$C_{(u,v)} : \{1,...,k\} \to \{1,...,k\}$, שהיא האילוץ שמשרה קודקוד המוצא של הצלע על קודקוד המטרה. אם u,v שני קודקודים בגרף הייצוג, ויש צלע (u,v) בגרף.

אזי עבור הקודקוד u ניצור סט משתנים מתאים V_u, שמאונדקס ע"י $\{-1,1\}^k$, באופן דומה V_v יתאים לקודקוד v ויהווה קבוצת משתנים מאונדקסת ע"י $\{-1,1\}^k$.

הערה 8.1 אנחנו עובדים עם עולם של $1,-1$, אבל בעצם עולם המשוואות הליניאריות מעל $\mathbb{Z}_2$ הוא עולם של $1,0$, כדי לתרגם נתבונן ב $\{1,-1\}$עם הכפל, במקום $(0,1)$ בשביל חיבור (זו אותה חבורה בעצם).

כעת לכל משתנה בבעיה המקורית יש לנו 2^k משתנים בבעיה החדשה, וההשמה מהמשתנה המקורי מגדירה באופן יחיד השמה ל 2^k המשתנים החדשים, כלומר יש $f_u : \{\pm 1\}^k \to \{\pm 1\}$ המתאימה לכל אחד מ 2^k המשתנים את ההשמה שלו.

כעת נתרגם את $C_{(u,v)}$ למשוואות ליניאריות על איברי V_u ו V_v.

כל משוואה (עבור $k=4$ למשל) תיראה כך:

$$f_u(1,1,-1,-1)\cdot f_v(-1,-1,-1,1)\cdot f_v(1,1,1,1)=-1$$

(כאן הכפל החליף את החיבור, אבל מכיוון שזו עדיין אותה חבורה $\mathbb{Z}_2$ זה שקול)

אנחנו יודעים לבנות את המשתנים (מסוף ההרצאה הקודמת) וכעת צריך להציג את האופן שבו נתרגם את $C_{(u,v)}$ למערכת משוואות ליניאריות בשלושה משתנים.

כעת מהשמה A לבעיה ב $LC(k)$, נוכל לייצר (על סמך בניית המשתנים) את A', השמה ל 2^k משתנים באופן שאם ההשמה לu היא i, אז $f_u(x)=x_i$ (כאן x הוא וקטור k מקומי שמאנדקס משתנה ב 2^k המשתנים שהם תמונת u ברדוקציה). כלומר f_u היא הקידוד ב"קוד ארוך" ($Long\,Code$) של i.

לדוגמה - עבור $k=3$:

$$A(4)=2$$

יש לנו את הוקטורים

$$(-1,-1,-1),(-1,-1,1),(-1,1,-1),(-1,1,1),...,(1,1,1)$$

וההשמה תהיה בחירת הביט השני מכולם, דהיינו:

$$(-1,-1,1,1,...,1)$$

וזה הוקטור ה 2^k מקומי המייצג את ההשמה A' לכל אחד מ 2^k המשתנים שהם תרגום u.

8.1 דרישת שלמות

אם $A(u)=i, A(v)=j$ ומתקיים ש $C_{(u,v)}(i)=j$ (כלומר ההשמה A מספקת את האילוץ $C_{(u,v)}$), אז A' צריכה לספק לפחות $1-\delta$ מ$C'_{(u,v)}$ (אוסף המשוואות הליניאריות שהן תירגום האילוץ $C_{(u,v)}$). המשפט האחרון מגלם את דרישת השלמות מהרדוקציה, כלומר שבאמת בעיות ספיקות לחלוטין יתורגמו לבעיות שהן לפחות $1-\delta$ ספיקות.

8.2 נסיון ראשון לדרישת נאותות:

אם A' היא השמה שהיא לפחות $\frac{1}{2}+\varepsilon'$ מספקת את $C'_{(u,v)}$, אז A' מהווה קידוד חוקי של השמה כלשהי A שנותנת ל u את i ול v את j כך ש $C_{(u,v)}(i)=j$. נעיר כי זו דרישה חזקה מאד שנובע ממנה שאם A' היא לא קידוד חוקי ב"קוד ארוך", אזי היא מספקת פחות מ$\frac{1}{2}+\varepsilon'$ מ $C'_{(u,v)}$.

אבל הדרישה הזו היא בלתי אפשרית, מדוע? נניח שיש השמה שמספקת את A, אזי מדרישת השלמות יש A' שמספקת $1-\delta$, ניקח את A', ונחליף את ההשמה לאחד מ 2^k המשתנים. נקבל שכרגע $1-\delta-\varepsilon''$ (כאשר ε'' קטן מאד) מהאילוצים מסופקים, אולם A' החדשה אינה עוד קידוד חוקי לפי "קוד ארוך".

8.3 נסיון שני לדרישת נאותות:

קיימת העתקה ψ שמקבלת פונקציה $f:\{\pm1\}^k\to\{\pm1\}$, ונותנת $\psi(f)\in\{1,...,k\}$ כך שאם f_u ו f_v מספקות יותר מ $\frac{1}{2}+\varepsilon'''$ מהמשוואות ב $C'_{(u,v)}$, אזי $\psi(f_u),\psi(f_v)$ מספקות את $C_{(u,v)}$. אבל כנראה שגם זו דרישה בעייתית מדי.

8.4 נסיון שלישי לדרישת נאותות:

ψ העתקה מפונקציות $f:\{\pm1\}^k\to\{\pm1\}$ ל $D=\psi(f)$, כך ש D התפלגות על $\{1,...,k\}$ שמקיימת: אם f_u,f_v מספקות יותר מ $\frac{1}{2}+\alpha$ מ $C'_{(u,v)}$, אז

$$Pr\left(C_{(u,v)}(i)=j\right)>\phi(\alpha)$$

$$i\sim\psi(f_u)$$
$$j\sim\psi(f_v)$$

(כאן ϕ היא סתם פונקציה כלשהי של α)

נוכיח כי הדרישה השלישית היא טובה:

נניח ש A' מספקת יותר מ $\frac{1}{2}+\varepsilon'$ מ $\bigcup_{(u,v)\in E}C'_{(u,v)}$.

אזי יש מספר קבוע של $C'_{(u,v)}$־ים (שמהווים נאמר החלק ה $\varepsilon^{(2)}$ מתוך כלל ה $C'_{(u,v)}$ים) אשר הם לפחות $\frac{1}{2}+\alpha$ ספיקים (אחרת הסכום הכולל לא היה יותר מחצי).

כעת לכל u, נבנה השמה $A(u)$ ע"י דגימה אקראית מתוך $\psi(f_u)$. לכל $C'_{(u,v)}$ מתוך המספר הקבוע של ה $C'_{(u,v)}$ים שהם לפחות $\frac{1}{2}+\alpha$ ספיקים, A מספקת את $C_{(u,v)}$ בהסתברות $\phi(\alpha)$. בתוחלת - יותר מ $\varepsilon^{(2)}\cdot\phi(\alpha)$ מהאילוצים $C_{(u,v)}$ מסתפקים ע"י A $\Leftarrow$כלומר בעיית ה LC המקורית שלנו היא לפחות $\varepsilon^{(2)}\cdot\phi(\alpha)$ ספיקה. כעת אם נבחר את ε המקורי להיות קטן מגודל זה, נקבל את הרדוקציה המבוקשת.

עד כאן הראינו שדרישה 3 אכן מספקת נאותות, אבל לא הצגנו שום דרך קונסטרוקטיבית לתרגם את האילוצים, ולא הצגנו שום ψ, אשר קיומה נדרש לנו.

8.5 בניית המשוואות הליניאריות

8.5.1 נסיון ראשון

ניקח $x,y\in_R\{-1,1\}^k$ (כלומר נבחר באקראי). המשוואה תהיה:

$$f_u(x)\cdot f_v(y)\cdot f_u(z)=1$$

אם f_u הוא קידוד של i וכנ"ל f_v הוא קידוד של j, אז משוואה זו היא בעצם:

$$x_i\cdot y_j\cdot z_i=1$$

נניח ש $j=C_{(u,v)}(i)$ ואז נוכל לכתוב:

$$x_i\cdot y_{C_{(u,v)}(i)}\cdot z_i=1$$

קל לוודא שכדי שזה יתקיים חייב להיות $z_i = x_i \cdot y_{C_{(u,v)}(i)}$, וכיוון שלא ידוע לנו i, זה צריך להיות נכון לכל i, כלומר המשוואה:

$$f_u(x) \cdot f_v(y) \cdot f_u(z) = 1$$

עבור z המוגדר על ידי $z_i = x_i \cdot y_{C_{(u,v)}(i)}$

8.5.2 נסיון שני

ניקח את z_i להיות כפי שהגדרנו אותו בנסיון הקודם: $z_i = x_i \cdot y_{C_{(u,v)}(i)}$ בהסתברות $1-\delta$, ו $z_i = -x_i \cdot y_{C_{(u,v)}(i)}$ בהסתברות δ.

9 :

נחזור להגדרת אופן בניית המשוואות מהשיעור הקודם.
נתבונן ב C', אוסף המשוואות הלינאריות עם מציין 2 שבנינו, בתור התפלגות על 2^k משתנים.
ניקח $x, y \sim \{-1.1\}^k$, ונרשום את המשוואה:

$$f_u(x) f_v(y) f_u(?) =$$

כעת כיוון ש $f_u(x) = x_i$ ו $f_v(y) = y_j$ נוכל לבחור את "?" להיות כזה שכשכופלים אותו ב $f_u(x) f_v(x)$, מקבלים

$$f_u(x) f_v(y) f_u(?) = 1$$

ברור שאם נבחר ? כך ש $f_u(?) = x_i y_j = x_i y_{C(i)}$ נקבל את המבוקש.
נסמן את $x_i y_{C(i)}$ ע"י $x \cdot y^c$
עם זאת - אם נבחר את f להיות קבועה על 1 כל המשוואות יסתפקו, אז אנחנו בבעיה.
אם היינו בוחרים את המשוואות להיות מהצורה $\ldots = -1$ אז הפונקציה הקבועה על -1 תספק את כל המשוואות.
הבעיה הזאת נפתרת על ידי הגרלה, כלומר המשוואות שניקח תהיינה:

$$f_u(x) f_v(y) f_u(\zeta \cdot x \cdot y^c) = \zeta$$

כאשר $\zeta \sim \{-1, 1\}$
אז נבחר $Z_1, \ldots, Z_k$ כך ש $Z_i = \begin{cases} 1 & w.p.\ 1-\varepsilon \\ -1 & w.p.\ \varepsilon \end{cases}$, ונחליף את המשוואה ל:

$$f_u(x) f_v(y) f_u(\zeta \cdot x \cdot y^c \cdot Z) = \zeta$$

מדוע עשינו זאת?
נשים לב שבמצב הקודם של המשוואה - לא רק השמות מספקות של f_u, f_v יספקו את המשוואה, אלא גם, למשל, מכפלות באורך 3 של השמות מספקות.
למשל: נניח ש u_1, v_1 השמה מספקת וכן u_2, v_2 וגם u_3, v_3, ואז נקבל כי עבור השמה ל 2^k המשתנים שלנו שהיא $f_{u_1} \cdot f_{u_2} \cdot f_{u_3}$ ו $f_{v_1} \cdot f_{v_2} \cdot f_{v_3}$, נקבל שתוצאת המשוואה תהיה ζ^3 אשר שווה לζ.
לכן הוספנו את אלמנט ה Z, ואז במקרה של מכפלות ארוכות של השמות חוקיות, נקבל כי

$$f_u(x) f_v(y) f_u(\zeta \cdot x \cdot y^c \cdot Z) = \zeta \cdot \prod_{i \in S} Z_i$$

והתוחלת של מכפלות של Z_i שואפת לאפס כאשר המכפלות הן ארוכות (S היא קבוצת האינדקסים במכפלה).

$$E\left[\prod Z_i\right] = \prod E[Z_i] = (1-2\varepsilon)^{|S|}$$

9.1 עוד הגדרות

הגדרה 9.1 נגדיר עבור $S \subseteq \{1,...,k\}$ את הפונקציה:

$$\chi_S(x) := \prod_{i \in S} x_i$$

כאשר כמקודם $x \in \{-1,1\}^k$

נתבונן באוסף הפונקציות מ $\{-1.1\}^k$ לממשיים, שיסומן ע"י $\mathbb{R}^{\{\pm 1\}^k}$, אשר קל לראות שמגדיר מרחב וקטורי ממימד 2^k מעל הממשיים.

הגדרה 9.2 מכפלה פנימית על המרחב לעיל ע"י:

$$\langle f, g \rangle = E_{x \sim \{\pm 1\}^k}[f(x)\, g(x)]$$

ואז:

$$\langle \chi_S, \chi_T \rangle = E\left[\prod_{i \in S} x_i \cdot \prod_{j \in T} x_j\right] = E\left[\prod_{i \in S \Delta T} x_i\right] = \prod_{i \in S \Delta T} E[x_i] = \begin{cases} 1 & S = T \\ 0 & S \neq T \end{cases}$$

ולכן הקבוצה $\{\chi_A : A \subseteq \{1,...,k\}\}$ היא בגודל 2^k וכן כל שני וקטורים בתוכה ניצבים זל"ז, וארכם 1, ולכן זהו בסיס אורתונורמלי למרחב הוקטורי.

הערה 9.3 לכל פונקציה $f : \{\pm 1\}^k \to \mathbb{R}$ קיימת דרך יחידה לרשום את f באופן הבא:

$$f = \sum_{S \subseteq \{1,...,k\}} \hat{f}(S) \cdot \chi_S$$

ונובע מכך שניתן לחלץ את המקדמים $\hat{f}(S)$ ע"י:

$$\hat{f}(S) = \langle f, \chi_S \rangle$$

כעת אם הטווח של f במקום הממשיים הוא $\{\pm 1\}$, אז נקבל:

$$\sum \hat{f}(S)^2 = \langle f, f \rangle = 1$$

9.2 קצת אנליזה

כעת בהנתן f_u, f_v , נרשום $f_u = \sum_S \hat{f}_u(S) \cdot \chi_S$ וגם $f_v = \sum_S \hat{f}_v(S) \cdot \chi_S$ וננתח את:

$$E\left[f_u(x)\, f_v(y)\, f_u\left(\zeta \cdot x \cdot y^C \cdot Z\right) \cdot \zeta\right] = Pr(sat) - Pr(unsat)$$

$$\begin{aligned} & E\left[\left(\sum_S \hat{f}_u(S)\, \chi_S(x)\right)\left(\sum_R \hat{f}_v(R)\, \chi_R(y)\right)\left(\sum_T \hat{f}_u(T)\, \chi_T\left(\zeta \cdot x \cdot y^C \cdot Z\right)\right) \cdot \zeta\right] \\ =& E\left[\sum_{S,R,T} \hat{f}_u(S)\, \hat{f}_v(R)\, \hat{f}_u(T) ...\right] \\ =& \sum_{S,R,T} \hat{f}_u(S)\, \hat{f}_v(R)\, \hat{f}_u(T) \cdot E\left[\chi_S(x)\, \chi_R(y)\, \chi_T\left(\left(\zeta \cdot x \cdot y^C \cdot Z\right)\right) \cdot \zeta\right] \end{aligned}$$

מכיוון שפונקציית ה χ כפי שהגדרנו אותה מקיימת $\chi_S(a\cdot b)=\chi_S(a)\chi_S(b)$, וכיוון ש ζ הוא קבוע (1 או -1) אז מתקיים:

$$\chi_T(\zeta\cdot x\cdot ...)=\prod_{i\in T}\zeta\cdot x_i\cdot ...$$

ולכן מהמכפלה הנ"ל מתקיים:

$$=\sum_{S,R,T}\hat{f}_u(S)\hat{f}_v(R)\hat{f}_u(T)\cdot E\left[\chi_S(x)\chi_R(y)\cdot\chi_T(x)\cdot\chi_T\left(y^C\right)\cdot\chi_T(Z)\cdot\zeta^{|T|}\cdot\zeta\right]$$
$$=\sum_{S,R,T}\hat{f}_u(S)\hat{f}_v(R)\hat{f}_u(T)\cdot E\left[\chi_S(x)\chi_R(y)\cdot\chi_T(x)\cdot\chi_T\left(y^C\right)\cdot\chi_T(Z)\right]\cdot\underbrace{E\left[\zeta^{|T|+1}\right]}_{0\,if\,T\,is\,even}$$

וזה נותן 0 אם $|T|$ זוגי, כי התוחלת של ζ מאפסת את הביטוי, אחרת - התוחלת של ζ תיתן 1 , ונתמקד בחלק האחר של התוחלת, ניזכר שה χ הן בסיס אורתונורמלי, ולכן מכפלת χ_S ו χ_T היא 0 אם $T\neq S$, ו 1 אחרת. כלומר הביטוי לא יתאפס רק אם $S=T$ וגם $|T|$אי זוגי.
נתבונן כעת בביטוי שמופיע:

$$\chi_T\left(y^C\right)=\prod_{i\in T}\left(y^C\right)_i=\prod_{i\in T}y_{C(i)}$$

נגדיר:

$$C_2(T)=\{j:\#\{i\in T:C(i)=j\}\ is\,odd\}$$

ואז

$$\prod_{i\in T}y_{C(i)}=\prod_{j\in C_2(T)}y_i=\chi_{C_2(T)}(y)$$

והביטוי הקודם שקיבלנו הוא:

$$=\sum_{S,R,T}\hat{f}_u(S)\hat{f}_v(R)\hat{f}_u(T)\cdot E\left[\chi_S(x)\underbrace{\chi_R(y)}\cdot\chi_T(x)\cdot\underbrace{\chi_{C_2(T)}(y)}\cdot\chi_T(Z)\right]\cdot\underbrace{E\left[\zeta^{|T|+1}\right]}_{0\,if\,T\,is\,even}$$

נשים לב שמנימוק דומה לקודם (האורתונורמליות) מתקבל $\chi_{C_2}(y)\cdot\chi_R(y)\neq 0$ רק אם $C_2(T)=R$.
לכן כדי שהביטוי לא יתאפס דרוש לנו בו זמנית $S=T$ וגם $C_2=R$ וגם $|T|$ אי־זוגי.
תחת ההנחות הללו הביטוי הנ"ל שווה ל:

$$\sum_{S:|S|\,is\,odd}\hat{f}_u(S)^2\hat{f}_v(C_2(S))\cdot(1-2\varepsilon)^{|S|}$$

10 המשך הרדוקציה למשוואות ליניאריות

וזה נותן את ההסתברות ש f_u,f_v מסתפקות פחות ההסתברות שאינן מסתפקות (כי התוחלת כופלת ב 1 משוואות מסתפקות וב -1 משוואות שאינן מסתפקות), או גם:

$$2Pr\left[f_u,f_v\,satisfy\,e\right]_{e\sim C'}-1$$

בסכום שלעיל יתקבל משהו שאינו אפס רק כאשר S היא היחידון i ו $C_2(S)$ היחידון j, ואז יתקבל בסכום $1-2\varepsilon$, כלומר אם f_u,f_v קידודים חוקיים בקידוד ארוך - אז נקבל את תכונה 1 שרצינו, שכן ההסתברות לספק את f_u,f_v גדולה או שווה ל $1-\varepsilon$ כפי שרצינו.

נרצה לקבל את תכונה 2: נניח שמתקיים (*)

$$Pr\left[f_u, f_v \, satisfies \, e\right]_{e\sim C'} \geq 1-\varepsilon-\alpha$$

נרצה כעת פונקציית פיענוח ψ מ f (פונקציות) ל $\{1,...,k\}$ כך שאם אכן מתקיים (*) אז יתקיים $C\left(\psi\left(f_u\right)\right)=\psi\left(f_v\right)$
ניקח עבור $\psi\left(f\right)$ להיות הקואורדינטה המינימלית i כך ש $\hat{f}\left(\{i\}\right)$ מקסימלי. כעת כיוון ש:

$$f\left(x\right)=X_i=1\cdot\chi_{\{i\}}\left(x\right)$$

$$\hat{f}\left(\{i\}\right)=1$$

כעת אם מתקיים $(*)$, מתכונה 1 מובטח לנו שהתוחלת היא פעמיים ההסתברות ש f_u, f_v יסתפקו פחות 1, ההסתברות הזו היא לפי $(*)$ לפחות $1-\varepsilon-\alpha$ ולכן התוחלת היא לפחות:

$$1-2\varepsilon-2\alpha$$

נסמן:

$$\nabla=\sum_{S:|S| \, is \, odd}\hat{f}_u\left(S\right)^2\hat{f}_v\left(C_2\left(S\right)\right)\cdot\left(1-2\varepsilon\right)^{|S|}$$

וכעת ידוע לנו כי:

$$1-2\varepsilon-2\alpha\leq\nabla\leq\left(1-2\varepsilon\right)\sum_{|S|=1}\hat{f}_u\left(S\right)^2f_v\left(C_2\left(S\right)\right)+\underbrace{\sum_{|S| \, odd \, and>1}\left(1-2\varepsilon\right)^{|S|}\hat{f}_u\left(S\right)^2f_v\left(C_2\left(S\right)\right)}_{A}$$

נשים לב כי ב A ידוע לנו ש $\hat{f}_v<1$ ולכן נוכל לוותר עליו ולכל היותר נגדיל את הביטוי, כמו כן כיוון ש $|S|\geq 3$ אז $\left(1-2\varepsilon\right)^{|S|}\leq\left(1-2\varepsilon\right)^3$ ולכן נחליף ב A בין שני אלו ונקבל:

$$\begin{aligned}\nabla&\leq\left(1-2\varepsilon\right)\sum_{|S|=1}\hat{f}_u\left(S\right)^2f_v\left(C_2\left(S\right)\right)+\sum_{|S| \, odd \, and>1}\left(1-2\varepsilon\right)^3\hat{f}_u\left(S\right)^2\\&\leq\left(1-2\varepsilon\right)\sum_{|S|=1}\hat{f}_u\left(S\right)^2f_v\left(C_2\left(S\right)\right)+\left(1-2\varepsilon\right)^3\sum_{|S| \, odd \, and>1}\hat{f}_u\left(S\right)^2\end{aligned}$$

כעת כמובן $\sum\hat{f}_u^2=1$ ולכן $\sum_{|S|\geq 3}\hat{f}_u^2\leq 1-\sum_{|S|=1}\hat{f}_u^2$, נציב זאת ונקבל:

$$\begin{aligned}&\leq\left(1-2\varepsilon\right)\sum_{|S|=1}\hat{f}_u\left(S\right)^2f_v\left(C_2\left(S\right)\right)+\left(1-2\varepsilon\right)^3\left(1-\sum_{|S|=1}\hat{f}_u\left(S\right)^2\right)\\&=\left(\sum_{|S|=1}\hat{f}_u\left(S\right)^2\right)\left(1-2\varepsilon-\left(1-2\varepsilon\right)^3\right)+\left(1-2\varepsilon\right)^3\end{aligned}$$

ולכן:

$$\sum_{|S|=1}\hat{f}_u\left(S\right)^2\geq\frac{1-2\varepsilon-2\alpha-\left(1-2\varepsilon\right)^3}{1-2\varepsilon-\left(1-2\varepsilon\right)^3}=1-O\left(\frac{\alpha}{\varepsilon}\right)$$

ולכן:

$$\sum_{|S|\geq 3}\hat{f}_u\left(S\right)^2 < O\left(\frac{\alpha}{\varepsilon}\right)$$

$$\begin{aligned}\sum_{|S|=1}\hat{f}_u\left(S\right)^2\hat{f}_v\left(C_2\left(S\right)\right)\cdot\left(1-2\varepsilon\right) &= \sum_i \hat{f}_u\left(\{i\}\right)^2\hat{f}_v\left(C\left(\{i\}\right)\right)\cdot\left(1-2\varepsilon\right)\\ &\geq 1-2\varepsilon-2\alpha-O\left(\frac{\alpha}{\varepsilon}\right)\end{aligned}$$

ולכן:

$$\begin{aligned}\sum_i \hat{f}_u\left(\{i\}\right)^2\hat{f}_v\left(C\left(\{i\}\right)\right) &\geq \frac{1-2\varepsilon-2\alpha-O\left(\frac{\alpha}{\varepsilon}\right)}{1-2\varepsilon}\\ &\geq 1-4\alpha-O\left(\frac{\alpha}{\varepsilon}\right)\end{aligned}$$

כיוון שהסכום של כל ה $\hat{f_u^2} < 1$ אפשר להסתכל עליהם כמשקלים בממוצע משוקלל על $\hat{f}_v$ ולכן קיבלנו שיש לנו ממוצע משוקלל של ה $\hat{f}_v$ -ים שגדול מגודל מסויים, מכך נובע - שיש לפחות $\hat{f}_v\left(C\left(i\right)\right)$ אחד, שהוא גדול או שווה לגודל המסויים הזה. כלומר:

$$\exists C\left(i\right)\ s.t.\ \hat{f}_v\left(C\left(i\right)\right)\geq 1-4\alpha-O\left(\frac{\alpha}{\varepsilon}\right)$$

כעת לא זאת בלבד שיש אחד שכזה, יש לכל היותר אחד, שכן α קטן מאד, ולכן גודל זה קרוב מאד ל 1, ולא יכולים להיות שניים.

מכאן נובע שבסכום (ה"ממוצע המשוקלל" שלנו) יש רכיב אחד (כלומר j אחד) שמהווה את רוב ערכו של הסכום, כלומר $\hat{f}_v\left(j\right)$ כך שגודל זה גדול מ $1-O\left(\alpha\right)$.

כלומר:

$$\sum_{i:C(i)=j}\hat{f}_u\left(i\right)^2\hat{f}_v\left(j\right)\geq 1-4\alpha-O\left(\frac{\alpha}{3}\right)$$

ואז:

$$\sum_{i:C(i)=j}\hat{f}_u\left(i\right)^2\geq 1-4\alpha-O\left(\frac{\alpha}{3}\right)$$

וממשפט של פרידגוט, קלעי ונאור אנחנו מקבלים שעבור פונקציה בוליאנית שקרובה מאד ל 1 - לא יתכן שהמקדמים "מתפזרים", כלומר חייב להיות מקדם שהוא הרבה יותר דומיננטי מכל האחרים, ואז נקבל שקיים ויחיד i שמקיים $C\left(i\right)=j$ כפי שרצינו.

אם נרצה לוותר על השימוש במשפט, נשנה את ψ באופן הבא:

$\psi\left(f\right)$: בחר i בהסתברות פרופורציונית ל $\hat{f}\left(i\right)^2$

כעת נרצה לקבל את תכונת הנאותות השלישית, ולכן נניח שמתקיים:

$$Pr\left[f_u, f_v\ satisfy\ e\right]_{e\sim C'}\geq\frac{1}{2}+\alpha$$

ואז נקבל, בדומה למקרה הקודם שמתקיים:

$$2\alpha\leq\nabla$$

נרצה ψ מ f ל $\{1,...,k\}$ כך שיתקיים $Pr\left[C\left(\psi\left(f_u\right)\right)=\psi\left(f_v\right)\right]\geq\phi\left(\alpha,\varepsilon\right)>0$

(כלומר מ f יתקבל וקטור באורך k שסכום הקואורדינטות שלו 1, שבכל קואורדינטה ההסתברות של האינדקס שלה)

ראינו קודם שכאשר ∇ קרובה ל 1, מרבית המשקל בסכום מתקבל מיחידונים (כלומר $|S|=1$), וכאן אנחנו לא יכולים לומר זאת.

יהי $l=\log_{1-2\varepsilon}\alpha$ ואז

$$\sum_{|S|\geq l}(...)\leq\alpha$$

וכיוון שהסכום על כל גדלי S הוא $\nabla=2\alpha$, נובע מכך כי:

$$\sum_{|S|<l\,and\,is\,odd}\hat{f}_u(S)^2\,\hat{f}_v(C_2(S))\,(1-2\varepsilon)^{|S|}\geq\alpha$$

כעת אפשר לוותר על ההכפלה ב $(1-2\varepsilon)$ ולשמור על אי השוויון (נזכור כי הכנסנו אותה במקור רק כי הכנסנו למערכת "רעש" באמצעות Z כדי למנוע ממכפלות ארוכות של f להשפיע על התוצאה הסופית, ואנחנו מתעסקים עם מכפלות באורך חסום כרגע) ולכן:

$$\sum_{|S|<l\,and\,is\,odd}\hat{f}_u(S)^2\,\hat{f}_v(C_2(S))\geq\alpha$$

כעת נוסיף לאילוץ הסכימה את האילוץ שאנחנו סוכמים רק על איברים שעבורם $\hat{f}_v(C_2(S))\geq\frac{\alpha}{2}$, מדוע זה בסדר? מכיוון שיש לנו ממוצע ממושקל שהוא גדול מ α, אזי בפרט אם נזרוק את כל האיברים שהם קטנים מ $\frac{\alpha}{2}$, לא נקטין את הסכום הכולל יותר מ $\frac{\alpha}{2}$, ולכן:

$$\sum_{|S|<l\,and\,is\,odd\,and\,\hat{f}_v(C_2(S))\geq\frac{\alpha}{2}}\hat{f}_u(S)^2\,\hat{f}_v(C_2(S))\geq\frac{\alpha}{2}$$

נגדיר את $\psi(f)$ כך שבסיכוי $\frac{1}{2}$:

[פיענוח לפי u] נבחר Sבסיכוי $\hat{f}_u(S)^2$ ואז נבחר $i\in S$ רנדומי וזה יהיה $\psi(f)$.

ובסיכוי $\frac{1}{2}$:

[פיענוח לפי v] נסתכל על Sים כך ש־ $\hat{f}(S)\geq\frac{\alpha}{2}$ ונבחר אחד מהם בסיכוי פרופורציוני ל $\hat{f}(S)$, נבחר $j\in S$ מקרי וזה $\psi(f)$. (אם אין $\hat{f}(S)$ לפי דרישתנו, נבחר $j=1$)

ננתח את הסיכוי ש ψ מספקת את C:

$$Pr[\psi\,satisfies\,C]\geq\frac{1}{4}\cdot\sum_{|S|<l\,and\,is\,odd\,and\,\hat{f}_v(C_2(S))\geq\frac{\alpha}{2}}Pr[We\,chose\,S\,for\,u\,and\,C_2(S)\;\;for\,v]\cdot\frac{1}{l}$$

כי בהנתן שבחרנו "נכון", אז אכן האילוץ מתקיים בהסתברות לפחות $\frac{1}{l}$.

וכיוון שבחרנו פרופורציונית, צריך לחשב מהו קבוע הנירמול שלנו, ידוע שסכום הריבועים של ה $\hat{f}_v$ הוא 1 ולכן:

$$\sum_{f_v(T)>\frac{\alpha}{2}}\hat{f}_v(T)\cdot1\underbrace{\leq}_{cauchy-shcwartz}1\cdot\sqrt{\frac{4}{\alpha^2}}\leq\frac{2}{\alpha}$$

ולכן:

$$\frac{\alpha}{8l}\cdot\sum_{|S|...}\hat{f}_u(S)^2\,\hat{f}_v(C_2(S))\geq\frac{\alpha^2}{16l}$$

נציב את הערך של l כפי שהגדרנו אותו ונקבל:

$$\approx\frac{\alpha^2\varepsilon}{8\log\left(\frac{1}{\alpha}\right)}$$

11 משפט ה PCP

הערה 11.1 אם במקום רדוקציה ל $E3-LIN$ ניקח רדוקציה ל $E3-SAT$ נוכל לקבל תמונה דומה, כאשר השמה מספקת ל u,v תיתן השמה מספקת ל V_u,V_v (עם האילוץ $C'_{(u,v)}$), ובמידה שיש השמה מספקת יותר מ $\frac{7}{8}+\alpha$ מקבוצות המשתנים בוליאניים, נוכל לקבל השמה שמספקת את u,v בהסתברות $\varepsilon(\alpha)$ כמקודם.

נרצה להראות כי בעיית הפער $\left(1-\varepsilon,\frac{7}{8}+\alpha\right)$ עבור $E3-SAT$ היא NP-קשה.
זה מתקבל בקלות מתוך ההוכחה שסיימנו בשיעור הקודם, מתחילים מבעיה ב $E3-LIN$, מתרגמים כל משוואה לארבע פסוקיות ב $E3-SAT$, כך למשל המשוואה:

$$x_{17}+x_{13}+x_7=1$$

מסתפקת ע"י 4 מתוך 8 ההשמות האפשריות למשתנים הללו, ולכן נגדיר 4 פסוקיות כך שכל אחת מהן לא מסתפקת על ידי אחת מההשמות ה"רעות", ולכן:

$$(x_{17}\vee x_{13}\vee x_7)$$

$$(\overline{x_{17}}\vee \overline{x_{13}}\vee x_7)$$

$$(x_{17}\vee \overline{x_{13}}\vee \overline{x_7})$$

$$(\overline{x_{17}}\vee x_{13}\vee \overline{x_7})$$

תהיה הרדוקציה של הנוסחה לעיל.
מכך נקבל שמה שהוא $1-\varepsilon$ ספיק יעבור לבעיה שאף היא $1-\varepsilon$ ספיקה. ומ $\frac{1}{2}+\alpha$ ספיקות נקבל $\frac{1}{2}+\alpha+\frac{3}{4}\left(\frac{1}{2}-\alpha\right)=\frac{7}{8}+\frac{\alpha}{4}$.

משפט 11.2 PCP- קיימים קבועים k ו $\varepsilon>0$ כך שבעיית הפער $(1,1-\varepsilon)$ עבור $2-csp(k)$ היא NP-קשה.
מכך נובע

$$NP\subseteq PCP^{na}_{1,1-\varepsilon}\left[O\left(\log n\right),2,\log k\right]$$

למה 11.3 ראשית - קיים k קבוע ו $\varepsilon_0>0$ קבוע, ואלגוריתם פולינומי כך שהקלט לאלגוריתם הוא בעיית $2-csp(k)$, נסמנה I, והפלט הוא בעיית $2-csp(k)$ שנסמנה I', כך ש:

1. $|I'|\leq C\cdot|I|$ (כלומר גודל הבעיה החדשה לא גדול מדי)
2. (שלמות) אם I ספיקה לחלוטין, כך גם I' (כלומר שלמות 1 נשמרת באלגוריתם).
3. (חיזוק) אם I הוא $1-\varepsilon$ ספיק, אז I' הוא $\max\{1-2\varepsilon,1-\varepsilon_0\}$ ספיק.

ומכך נובע משפט ה PCP ע"י הפעלת הלמה $\log n$ פעמים.

הגדרה קטנה: אם I היא $1-\varepsilon$ ספיקה, נאמר ש ε היא הגריעות ($badness$) של I.

11.1 האלגוריתם בקווים כלליים

11.1.1 שלב ראשון - יפוי ($beautifying$)

נייצר מ I את I^1, שהוא גם כן בעיית $2csp(k)$, המיוצגת ע"י גרף, שהוא גם גרף מרחיב. כמו כן - אם I ספיקה לחלוטין, כך גם I^1.
וכן:

$$Badness\left(I^1\right) \geq C_1 \cdot Badness\left(I\right)$$

(קיימנו את דרישת השלמות בלמה)

11.1.2 שלב שני - חיזוק ($amplifying$)

נייצר מ I^1 את I^2 תוך שמירה על קריטריון השלמות, כך שאו:

$$Badness\left(I^2\right) \geq C_2 \cdot \sqrt{t} \cdot Badness\left(I^1\right)$$

או:

$$Badness\left(I^2\right) \geq 1000 \cdot \varepsilon_0$$

כמו כן I^2 תהיה בעיה שהאלפבית של האילוצים שלה לא יהיה k אלא $k' = k'(k,t)$ (כלומר k' תלוי ב k,t) וגם k' יהיה גדול מ k בהרבה.

11.1.3 שלב שלישי - קידוד ($encoding$)

נייצר מ I^2 את I' תוך שמירה על קריטריון השלמות, כך ש:

$$Badness\left(I'\right) \geq C_e \cdot Badness\left(I^2\right)$$

ו I' יהיה מעל אלפבית בגודל k.

11.2 שלב היפוי

1. לייצר גרף בעל דרגה קבועה
2. לקבל הרחבה (כלומר גרף מרחיב מהגרף הקודם) בעלת דרגה קבועה

נתחיל דווקא מהשלב השני - יש בידינו גרף אילוצים בעל דרגה קבועה, ונרצה להרחיב אותו לכלל גרף מרחיב.

11.2.1 השגת דרגה קבועה

נחליף כל קודקוד u בקבוצת קודקודים u' שגדלה כדרגת u, ולכל צלע u,v תהיה צלע בין קודקוד יחיד ב u' לקודקוד יחיד ב v'.
למעשה החלפנו כל קודקוד ב"ענן" של קודקודים, שכל אחד מהם דרגתו 1. בתוך כל ענן u' - נבנה גרף מרחיב בין הקודקודים המשתתפים בענן, וכך נעשה לכל העננים, כך שדרגת כל הגרפים המרחיבים הללו קבועה d, וזה ניתן להשגה יחסית בקלות. הצלעות שבתוך כל ענן יהיו אילוצי שוויון, נקבל שהשמה מספקת ל I תספק גם את הגרף החדש שבנינו (אם נציב לכל הקודקודים ב u' את הערך שקיבל u קודם לכן). אפשר להראות שמתקיימת דרישת הנאותות, כיוון ש"שיפור" ההסתפקות של הגרף החדש יכול לנבוע רק מהשמת ערכים שונים לקודקודים בתוך ענן (אחרת זה שקול להשמה ל I), וניתן להראות שאילוצי השוויון על הקודקודים בגרף המרחיב שבענן, והיותו גרף מרחיב, גוררים כי על כל "רמאות" שכזו, שמאפשרת סיפוק צלע שלא היתה מסופקת על ידי השמה מיטבית בגרף המקורי, נשלם בצלעות שלא יסתפקו בתוך הענן עצמו, ולכן לא נצליח לשפר את ההסתפקות של הגרף החדש על ידי "רמאות" כגון זו, ולכן הנאותות תשתמר.

11.3 הרחבה וחיזוק

11.3.1 הרחבת הגרף

אם ניקח גרף מרחיב כלשהו על אותם קודקודים, ונוסיף את כל צלעותיו לגרף הקיים (במקום שכבר ישנן צלעות נוסיף צלעות כפולות)- נקבל גרף מרחיב (קל לוודא זאת), ואם הגרף המרחיב וגם הגרף שהתחלנו איתו היו רגולריים, נקבל גרף רגולרי מחיבורם. את הצלעות שנוסיף נגדיר להיות אילוצים טריוויאליים, כלומר כאלו המסתפקים תמיד, ולכן אנחנו מגדילים את הב $Badness$ בפקטור התלוי בדרגת הרגולריות של הגרף המרחיב, אבל זה קבוע ולכן:

$$new\,Badness \geq C_1 \cdot Old\,Badness$$

נתבונן בהילוך עצל (כלומר שבכל שלב יש גם הסתברות להשאר במקום) על הגרף, באורך t.

הגדרת התפלגות על המסלולים העצלים באורך $2t$: מתחילים בקודקוד אקראי, ואז מבצעים הילוך עצל באורך $2t$, מתקבלת התפלגות.
באופן שקול אפשר לבחור קודקוד אקראי x, לבצע הילוך עצל באורך t ועוד הילוך באורך t שמתחיל גם הוא ב x - ואז זה מגדיר התפלגות על מסלולים בין קודקודי המטרה (שהם לכל היותר באורך $2t$) .

שינוי התיוגים שאפשר לשים על קודקוד: נגדיר כעת שהשמה לקודקוד מכילה גם מידע על ההשמה לכל הקודקודים שהם במרחק של לכל היותר $2t$ ממנו, כלומר ההשמות הופכות להיות (אולי) מחרוזות ענקיות שרק תחילתן היא השמה מתוך $\{1,..,k\}$ לקודקוד, והמשכן הוא השמה לכל הקודקודים ברדיוס $2t$ מ v.

הגרף החדש שנבנה: בין כל v,u תהיה צלע רק אם יש ביניהם מסלול באורך קטן מ $2t$. האילוץ יהיה שההשמות לשניהם יסכימו על כל הקודקודים בחיתוך, והמשקל של הצלע יהיה ההסתברות למסלול זה לפי ההתפלגות שהגדרנו שנגזרת מהמסלולים העצלים.
נותר להראות שכעת אכן הגריעות עלתה, ועלות הניפוח קבועה.
ניפוח בקבוע - מספר האילוצים כעת $k^{d^{2t}}$ אבל זה קבוע, וכן כל עלויות הניפוח שלנו הן פונקציות של k,t ו d, כפול גודל הייצוג הקודם, וזה אכן מקיים את דרישתנו ש:

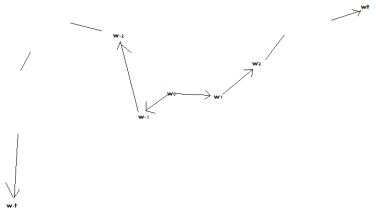
$$Badness\left(I^2\right) \geq C_2 \cdot \sqrt{t} \cdot Badness\left(I^1\right)$$

שלמות - נניח שיש השמה A' לגרף החדש שבנינו, ונרצה איזשהו קידוד שמייצר מ A' את A, השמה לגרף הישן. נראה ש A לא מספק חלק מסויים, ונוכיח מתוך כך שנובע ש A' לא מספק קבוע כפול אותו חלק מתוך הגרף החדש.

קידוד בהנתן A', ניקח קודקוד w כלשהו, נשים לב של w עצמו יש "דעה" לגבי ההשמה המתאימה לו ב A, וכך גם לכל הקודקודים שהם במרחק $2t$ ממנו בגרף הישן. נתבונן בכל הקודקודים שהם במרחק כזה מ w בגרף הישן, למשל עבור הקודקוד v - יש הסתברות למסלול w,v ודעתו של v משקלה הוא כהסתברות המסלול. מסכמים את כל המשקלים לדעה ש w צריך לקבל את הערך 1, את כל המשקלים לדעה שw צריך לקבל את הערך 2 , וכן הלאה. הדעה שיש לה הכי הרבה "תמיכה" - קובעת מה יהיה באמת משקלו של w בהשמה A שאנו בונים.
כעת כמובן

$$Badness_{I^1}\left(A\right) \geq Badness\left(I^1\right)$$

כיוון שהגריעות של I^1 היא מינימום על כל ההשמות שאפשר לבחור לו ו A היא השמה מסויימת, אז החלק היחסי שמסתפק ב A הוא לכל היותר החלק היחסי שניתן לספק על ידי השמה כלשהי.
נבחין כי לכל קודקוד w, וקודקוד v שמרוחק ממנו t צעדים - ההסתברות שהם מסכימים על ההשמה ל w היא לפחות $\frac{1}{k}$ (כי יש k אפשרויות לבחירת השמה, ונבחרה ההשמה הכי פופולארית).
נבחר הילוך עצל באורך $2t$ באופן אקראי (בגרף הישן), כזכור - באופן שקול ניתן לבחור קודקוד אקראי w_0 ושני מסלולים עצלים באורך t שמתחילים בו:



E_1= הצעד הראשון בהילוך הראשון מ w אינו עצל, וגם הצלע שהלכו עליה (שהיא בגרף הישן) אינה מסתפקת על ידי A. וגם w_tמסכים עם $A(w_1)$ ו w_{-t} מסכים עם $A(w_0)$.

עצם האפשרות שהצעד הראשון לא היה עצל היא בהסתברות חצי, וההסתברות לאי־הסתפקות צלע כלשהי היא $Badness_{I^1}(A)$.

ההסתברות שקודקודים במרחק t יסכימו על ההשמה של אחד מהם היא $\frac{1}{k}$, ולכן ההסתברות לשתי הסכמות בלתי תלויות כנ"ל היא לפחות $\frac{1}{k^2}$, ומכיוון שאחד המסלולים הוא באורך $t-1$ זה משנה בפקטור קבוע את ההסתברות ולכן היא $\frac{c}{k^2}$.

המאורעות הנ"ל בלתי תלויים ולכן נכפול את הסתברויותיהם.

$$Pr[E_1] \geq \frac{1}{2} \cdot \frac{c}{k^2} \cdot Badness_{I^1}(A) \geq \frac{c'}{k^2} \cdot B\left(I^1\right)$$

כעת נגדיר מאורע נוסף:

E_2=הצלע w_1, w_2 קיימת (כלומר הצעד הראשון בהילוך הראשון לא היה עצל), היא אינה מסתפקת ע"י A וגם w_t מסכים עם $A(w_2)$ ו w_{-t} מסכים עם $A(w_1)$.

ונגדיר את E_i לכל $\frac{-\sqrt{t}}{100} < i < \frac{\sqrt{t}}{100}$ ונקבל באופן דומה כי

$$Pr[E_i] \geq \frac{c''}{k^2} Badness\left(I^1\right)$$

ואז תוחלת מספר ה־i־ים כך ש E_i מתרחש היא:

$$\mathbb{E}[\#\{i : E_i\, occured\}] = \sum_i Pr[E_i] \geq \sqrt{t}\frac{c''}{k^2} \cdot Badness\left(I^1\right)$$

כלומר נקבל בממוצע מספר גדול של צלעות שאינן מסתפקות במסלול, נרצה להראות שזה לא נובע מכך שההתפלגות היא עם זנב כבד, כלומר המצב הוא לא שבהסתברות קטנה יש הרבה מאד מאד צלעות שאינן מסתפקות במסלול, אלא שבאופן כללי זהו ביטוי לכך ש"בדרך כלל" יש צלע אחת שלא מסתפקת.

זה נובע מכך שהגרף הוא גרף מרחיב (משלב הייפוי), ואז ההסתברות לראות n צלעות "רעות" (שאינן מסתפקות, ולכן הן שייכות לקבוצה קטנה יחסית) במסלול שאינו ארוך מדי ־ היא קטנה מאד.

11.4 קידוד

נרצה לייצר גרף מתאים כך שגודל האלפבית הוא k (כזכור בשלב הקודם הגדלנו אותו למשהו שהוא אקספוננציאלי ב k).

נוכל לקבל את זה באופן דומה לרדוקציה שעשינו ל $E3-LIN$ בעבר, אם נייצר תחילה גרף שהוא בעיית היטלים (כיסוי תוויות), וזאת נוכל לקבל אם לכל אילוץ נוסיף קודקוד ביניים שיוצר אילוצים על הקודקודים של אותו אילוץ, ואז נמשיך כמו ברדוקציה ההיא. כזכור בעבר איבדנו שלמות בתהליך, אבל אם נעשה את הרדוקציה ל $4-CNF$במקום ל $E3-LIN$, נוכל לשמור על השלמות.

נבצע עוד רדוקציה ל $2-csp$ (עשינו דבר דומה באחד מתרגילי העבר) וכך נקבל שוב גרף אילוצים כמו זה שהתחלנו בו, כאשר חיזקנו את הנאותות, כפי שרצינו.

12 תוספות ונספחים

$E4-CNF-SAT$ 12.1

נזכיר כי עבור הרדוקציה ל $E3-LIN$ הגדרנו f_u שאם היא קידוד חוקי לקוד ארוך אזי היא מקודדת השמה ל u, כנ"ל f_v.

כמו כן C הוא אילוץ $C_{u,v}:[k]\to[k]$, וממנו הגדרנו C' שהוא אוסף משוואות לינאריות בשלושה נעלמים (מעל $\mathbb{Z}_2$ ולכן אפשר כפל במקום חיבור):

$$f_u(x)\,f_v(y)\,f_u(x\cdot y^c)=1$$

הרעיון הכללי הוא לתרגם צומת ל"ענף" של צמתים, אילוץ לאוסף אילוצים ופונקציית פיענוח שמפענחת השמות מעננים שמסתפקים ע"י אוסף אילוצים להשמות למשתנים המקוריים שמספקות את האילוץ בהסתברות גבוהה.

בעבר פגענו במהלך העבודה בשלמות, ע"י הכנסת רעשים, הפעם נרצה לשמר אותה.

נבנה את הרדוקציה הפעם לא למשוואות אלא לבעיות $E4-SAT$. נגדיר לשם נוחות $FALSE\to 1$ ו $TRUE\to -1$.

ולכן

$$OR(1,-1)=-1$$
$$OR(1,1)=1$$

$$OR(\alpha,\beta)=-\frac{1}{2}-\frac{1}{2}\alpha-\frac{1}{2}\beta+\frac{1}{2}\alpha\beta$$

$$f_u(w)\vee f_u(x)\vee f_v(y)\vee f_u(z)$$

כאשר נגדיר ש:

$$z_i=\begin{cases}Random\,bit & w_i=1\\ -x_iy_{c(j)}w_1 & w_i=-1\end{cases}$$

יש הוכחה של זה במאמר של $Hastad$ שנקרא: $Some\,optimal\,inapproximability\,results$

UGC 12.2

$Unique\,Label\,Cover$ - ULC 12.2.1

כאשר האילוצים בין כל שני קודקודים הם חח"ע (כלומר כל צלע בגרף היא פרמוטציה).

הוכחנו בתרגיל ש $ULC\in P$

$UGC-Unique\,Game\,Conj.$ 12.2.2:

בעיית הפער $(1-\varepsilon,\delta)$ עבור $ULC(k)$, היא NP-שלמה (השערה של $Subhash\,Khot$)

כיסוי קודקודים - נתון גרף, ונרצה לבחור קבוצה מינימלית של קודקודים שמכסים את כל הקשתות (אפשר גם עבור היפר־גרף).

הוכח כי קשה לקרב את הפתרון עבור יותר מ $1.36...$ (ספרא ודינור).

סובאש ק'וט הוכיח קושי עבור $Min-2-UNSAT$ (בעיה עבור זוגות של ביטים), והוא הצליח לעשות זאת בהנחת UGC.

רגב וק'וט הראו שלקרב את בעיית כיסוי הקודקודים יותר מקירוב $2-\varepsilon$ (לכל ε) זה NP - קשה בהנחת UGC, וכמובן קירוב 2 יש לנו מאלגוריתם טריוויאלי. וזו כבר תוצאה די דרמטית לגבי בעיה שהיא מאד נחקרת.

$Max\,cut$ מציאת חתך מקסימלי בגרף.
נחשוב על הגרף כאילו צובעים את קודקדיו ב 1 או ב -1. ונרצה למצוא חיתוך בגרף כך שמקסימום צלעות מחברות -1 ו 1.
בעצם זה שקול לבעיית $E2-LIN$ מעל $\mathbb{Z}_2$, כי כל קשת מגדירה משוואה לינארית בשני נעלמים (ולכן זה NP שלם).
לקבל קירוב $\frac{1}{2}$ לבעיה הזו זה קל, אם בוחרים צלעות באקראי, וזו כבר תוצאה משנות השישים.
הצליחו לשפר זאת ל $\frac{1}{2}+|E|$ ודומים, כאשר התוספת לקירוב חצי תלויה במספר הקודקודים או הצלעות, ולכן אסימפטוטית זה לא עוזר במיוחד.
ב 94 הראו גומנס ווילאמסון ($Geomans\&Williamson$) אלגוריתם שמקרב לפחות עד כדי $0.8789...$, מאוחר יותר מצאו גרפים ספציפיים שעליהם לא מתקבל קירוב טוב יותר.
$KKMO$ (גיא ואחרים) הראו שהתוצאה הזו היא אופטימלית (כלומר קשה לקרב יותר מכך) בהנחת UGC.
ותוצאה של CMM (צ'אריקאם, מקריצ'ב ומקריצ'ב) מציגים אלגוריתם לבעיית ULC כללית, ולא רק בשתי תויות, כמו ש $Max-cut$, אלא לכל מספר של תוויות.
בעצם התקבל שאם UGC נכונה אז אנחנו יודעים בדיוק מה הם ε ו δ מניסוחו.
אם מישהו ימצא אלגוריתם טוב יותר מ CMM - הרי שהפרכנו את UGC. והרבה אנשים עובדים מצד אחד בכיוון הזה, ומאידך בנסיון להוכיח את ההשערה.
יש תוצאה די דרמטית של ראגאוואנדרה מהשנה שעברה (2008) שלכל בעיית אילוצים ($3-SAT$, $4-SAT$, $3-LIN$, כיסוי קודקודים ועוד ועוד) יש אלגוריתם פתרון שמתבסס על SDP ($Semi\,definite\,programming$), שזוהי בעצם הגישה הגיאומטרית של גומנס ושל ווילאמסון, שמשיג קירוב כפלי, ובהנחת UGC הקירוב הזה הוא אופטימלי. (למצוא את הקבוע ממש זו בעיה בסיבוכיות מאד גבוהה)

13 בעיות פתוחות וכיוונים להמשך

13.1 צמצום כמות המידע

נתעניין כרגע בכמות המידע שנמצאת בידי המוודא, לאחר שהמוכיח שכנע אותו בנכונות ההוכחה.
נרצה למזער כמות זו.
ברור שכאשר יש q שאילתות של ביט כ"א, מתקבל פרדיקט של $q-csp\,(2)$, ולכן השמה רנדומית תיתן נאותות שהיא לכל הפחות 2^{-q}. מדוע? כיוון שיש השמה שמספקת, ויש 2^q השמות, אז בתוחלת השמה רנדומית מסתפקות בחלקה ה $\frac{1}{2^q}$.
ומכאן

$$\log\frac{1}{s}\leq q$$

זאת ה $Amortized\,query\,complexity$ מאידך - לא ברור למה q הוא דווקא מדד טוב לכמות האינפורמציה שבידי המוודא.
שכן אם ההוכחה נכונה - נרצה לקבל מה שפחות מידע, אבל אם היא שגויה אז לא אכפת לנו, כי אין בזה נזק כביכול. למשל אם נקבל מידע עבור נוסחה של $3-LIN$:

$$x_1+x_{17}+x_{65}=1$$

אז בהנתן x_1 ו x_{17} יש רק השמה אחת ל x_{65} שעשויה לעניין אותנו - זאת שמספקת את הנוסחה, ובמקרה כזה ההשמה לא "תפתיע" אותנו. לכן לכאורה האינפורמציה הגלומה במשוואה לעיל היא 2 ביטים בלבד.
בהנתן פרדיקט כלשהו, אשר יש לו k השמות מספקות, בדומה למקרה הקודם, אנחנו מניחים שההשמה מסתפקת, ואז רק מעניין אותנו מהי ההשמה שמספקת אותו, וכדי לדעת את זה מספיק לקבל את האינדקס של ההשמה המספקת, כלומר $\log_2 k$ ביטים.
ולכן ה $Amortized\,free\,bit\,complexity$:

$$\frac{\log_2 k}{\log\frac{1}{s}}$$

אפשר להראות (ולא נראה) ש:

$$NP \subseteq PCP_{1,s}\left[free\,bit\,complexity\ \log k\right]$$

כך ש:

$$\frac{\log_2 k}{\log \frac{1}{s}} \leq \delta$$

לכל $\delta > 0$.

כלומר אחרי שקראנו 100 ביטים חופשיים, אפשר לקרוא עוד 10000 ביטים שאינם חופשיים (כלומר שאנחנו יודעים מראש מה ההשמה שהם צריכים לקבל כדי להסתפק), ולהקטין את הנאותות ל 2^{-10000}. כלומר להקטין את הנאותות בעזרת ביטים שאינם חופשיים, כלומר כאשר המוודא מקבל 0 אינפורמציה.

13.2 $Max-Cut$

נשים לב שה $a.f.b.c$ של $Max-Cut$ הוא 1 , שכן אם כל צלע היא אילוץ על שני קודקודים, שמספר ההשמות המספקות אותה הוא 2 (כשאחד הקודקודים צבוע 1 והשני צבוע -1).

13.3 המקרה הממוצע לעומת המקרה הגרוע ביותר

בדרך כלל אנחנו נוטים לנתח בעיות לפי הסיבוכיות הגרועה ביותר, אלא שבמקרה המעשי על פי רוב בעיה תהיה "ממוצעת", ולא מתוכננת כך שתהיה קשה ביותר.

נניח שיש לנו פסוקיות שהן בעיית $3-SAT$ שנגריל באקראי, כלומר קודם נבנה פסוקיות, אחר כך נשים בתוכן x־ים שלחלקם שלילה וחלקם לא, ואחר כך נחלק אינדקסים באקראי...

האם זו בעיה שהיא קלה בהסתברות גבוהה?

למשל עבור בעיית $Bounded-3-SAT$ - אם הקבוע שחוסם את מספר ההופעות של משתנה הוא קטן, אז בהסתברות טובה זו תהיה בעיה קלה, משום שכביכול האילוצים ישרו מעגלים קטנים שניתן יהיה לפתור בנפרד.

אם נפעל בדרך הפוכה - כלומר נבחר את מספר המשתנים, ואז מכל הפסוקיות האפשריות על משתנים אלו נבחר כל פסוקית בהסתברות קבועה - אזי אם ההסתברות קרובה ל־0 אז הביטוי כולו יהיה ספיק בהסתברות גבוהה, ולהיפך - אם ההסתברות לכל פסוקית היא גבוהה - אז בסבירות גבוהה נקבל הרבה סתירות ולכן זה לא יהיה ספיק. איפהשהו באמצע יש קבוע שאם נגריל לפיו (כלומר הוא יהיה ההסתברות לכל פסוקית) - בהסתברות חצי נקבל בעיה ספיקה, נכנה זאת P_c. קצת לפני P_c - יש קבוע שאם נגדיר אותו בתור ההסתברות לכל פסוקית - נקבל בעיה ספיקה בהסתברות מאד גבוהה. האם יש אלגוריתם פולינומיאלי למצוא לו השמה שכזו?

האם $3-SAT$ היא בעיה קשה בממוצע? בעיה פתוחה.

יש תחום שבו כן הראו קושי של בעיות בממוצע. בעיות סריג:

נניח יש $\{b_1,...,b_n\} \subseteq \mathbb{R}^n$, שהם בסיס. ואז $L = Span_{\mathbb{Z}_n}\left(\{b_1,...,b_n\}\right)$ הוא סריג.

נשאלת השאלה - מהי הנורמה האוקלידית של הוקטור הכי קצר בסריג? $Shortest\,vector\,problem$ = SVP

יודעים להראות שבעיה זו היא NP שלמה.

מסתבר שיש רדוקציה מהבעיה במקרה הגרוע ביותר, לבעיה במקרה הממוצע: בהנתן קירוב בפקטור כפלי $\sqrt{n}$ במקרה הממוצע - ניתן לקבל ממנו קירוב בפקטור כפלי n^5 עבור המקרה הגרוע ביותר. זו בעצם אחת התוצאות הכי טובות שקיימות לרדוקציה מהמקרה הממוצע למקרה הגרוע ביותר, וגם זה בעייתי, כיוון שרק ידוע לנו שלקרב את המקרה הגרוע ביותר עד כדי $2^{\sqrt{\log n}}$ זה NP־קשה, וזה קטן מכל פולינום. במילים אחרות - יש רדוקציה, אבל לא יודעים שמה שעושים רדוקציה אליו הוא באמת קשה.

13.4 UGC

נניח שנגריל בעיית ULC. מסתבר שלא ידועה לנו התפלגות שאם נגריל לפיה זה יהיה לא טריוויאלי לפתור.

בין השאר, משום שכשבוחרים צלעות באקראי מקבלים גרף מרחיב, ועבור גרפים מסוג זה, בעיית ה ULC היא קלה (כלומר יש לנו אלגוריתם פולינומיאלי).

13.4.1 (למה פיסיקאים מתעניינים בדברים כאלה?)

ב $3-SAT$ אפשר לדבר על שיעור הפסוקיות המסתפקות. בגביש מגנטי יש כל מיני אלקטרונים חופשיים עם ספין שהוא איזהשהו וקטור במרחב. אם רוצים לפשט את זה, אז הספינים מסתדרים או למטה או למעלה. יש אינטרקציות בין הספינים שאפשר לחשוב עליהן בתור אילוצים, וכאשר מקררים את הגביש הוא רוצה לשקוע לרמת אנרגיה נמוכה שבה מה שיותר אילוצים יסתפקו. הרבה מהאלגוריתמים לפתרון של בעיות $3-SAT$ מהצד של הסקאלה שבו כל פסוקית מוגרלת בהסתברות קטנה יחסית, מגיעים מהעולם הפיסיקאלי. חשבו גם לקחת בעיית $3-SAT$, להמיר אותה לבעיה פיזית על גביש כנ"ל, להניח את הגביש במקרר ולקבל פתרון.

זה לא עובד כיוון שהבעיה נתקעת במינימום לוקאלי. נעשו נסיונות לאלגוריתמים שמחממים את הגביש ומקררים שוב, וזה קצת יותר טוב, אבל בכל זאת מסתבר שלמרות התחזיות - הטבע לא נוטה באמת לרמת האנרגיה הנמוכה ביותר.