נושאים בתיאוריה של מדעי המחשב

על פי הרצאות מאת פרופ' נתי ליניאל וד"ר גיא קינדלר

9 ביוני 2010

רשם: שיר פלד, באמצעות LyX גרסה 1.6.1
תיקונים יתקבלו בברכה במהלך ההפסקות או בכתובת מייל $shirpeled@cs$
ספר שישמש אותנו לחלק מהנושאים $Computational\,Complexity\,:\,A\,Modern\,Approach$ מאת בועז ברק וסנג'יב ארורה.

1 שיעור 1

1.1 מבוא

1.1.1 פסאודו רנדומיות

כידוע יש בעיות הכרעה שאינן ב P, אולם קשה מאד להצביע על בעיה קונקרטית שכזו.
בעיה לדוגמא: האם קיים אלגוריתם A אשר בהנתן $n\in\mathbb{N}$, רץ בזמן $2^{n^{O(1)}}$ והפלט שלו הוא טבלת האמת של פונקציה $f_n:\{0,1\}^n\to\{0,1\}$ כך שהשפה

$$L=\left\{x:f_{|x|}\left(x\right)=1\right\}$$

אינה במחלקה P. (L היא בעצם השפה של f_n)
התשובה לכך אינה ידועה...

הגדרה 1.1 מספר נורמלי בבסיס b בקטע $[0,1]$ הוא מספר שבו כל מחרוזת חוזרת בהסתברות אחידה (ביחס לארכה). מספר נורמלי אם הוא נורמלי בכל בסיס.

לא ידועה לנו דרך לבנות מספר נורמלי באופן מפורש, ואף לא ידוע לנו על מספר שהוא נורמלי, אע"פ שבבחירת מספר אקראי בקטע - ההסתברות שהוא נורמלי היא 1 .
יש מבנים רבים שהם מסוג זה - כלומר שקשה לבנות אותם מפורשות אולם הם במובן מסויים מתקבלים בהסתברות גבוהה בבחירה אקראית. בין מבנים אלו אקספנדרים (גרפים מרחיבים), אקסטרקטורים, מחוללי פסאודו־רנדומיות.

1.1.2 כוחה של אקראיות

יש מודלים שבהם אי אפשר להחליף רנדומיות.
למשל - בתקשורת. לאליס ולבוב (A ו B) יש מחרוזות $x\in\{0,1\}^n$ ו $y\in\{0,1\}^n$ בהתאמה, ורוצים להכריע את בעיית הזהות. כלומר לחשב את:

$$f\left(x,y\right)=\begin{cases}1 & x=y\\ 0 & x\neq y\end{cases}$$

כמה ביטים נצטרך להעביר מעל גבי קו התקשורת?
קל להתרשם שניתן להכריע את הבעיה ע"י העברת n ביטים (כלומר כל הביטים של x לבוב, למשל).
מעניין גם שאי אפשר להסתפק בפחות מ n ביטים.
הוכחה סכמטית: אם אליס מעבירה ביט אחד לבוב, אזי יש x כך שמבחינת בוב מרחב האפשרויות הצטמצם בחצי לכל היותר, באינדוקציה - יוצא שצריך n ביטים כדי לכסות את כל 2^n המחרוזות האפשריות עבור x.

רנדומיות משותפת מה התשובה כאשר:

1. מוסיפים מקור משותף של ביטים מקריים

2. מסתפקים בכך שלכל x, y ההסתברות שהפרוטוקול נותן תשובה נכונה גדולה ממש מ $\frac{3}{4}$.

פתרון (כאשר נניח שבוב מנסה להעביר מידע לאליס שצריכה להחזיר את $f(x,y)$):
בוב ישדר

$$\left(\sum_{i=1}^{n} r_i \cdot y_i\right)(mod\,2)$$

ונחזור על זה פעמיים.
במילים אחרות - המקור הרנדומי (r) ישרה תת קבוצה של n, ונשלח את ביט הזוגיות עבור קבוצה זו. נעשה זאת פעמיים ולכן סה"כ נשלח 2 ביטים.
אם $x = y$ - ברור שהפרוטוקול תמיד יתן תשובה נכונה, כלומר בהסתברות 1, שזה אפילו יותר טוב ממה שדרשנו.
במקרה ש $x \neq y$ - אז יש i_0 כך ש $x_{i_0} \neq y_{i_0}$. נתבונן במחרוזת האקראית הראשונה שקיבלנו $r_1 r_2 ... r_{i_0} ... r_n$.
אחרי שקבענו את $\{r_i\}_{i \neq i_0}$ נתבונן בשני מקרים
במקרה הראשון:

$$\sum_{i \neq i_0} r_i y_i = \sum_{i \neq i_0} r_i x_i$$

במקרה זה - בהסתברות חצי $r_{i_0} = 1$ ובמקרה זה נקבל:

$$\sum_{i} r_i y_i \neq \sum_{i} r_i x_i$$

ולכן אליס תשיב נכונה.
במקרה השני:

$$\sum_{i \neq i_0} r_i y_i \neq \sum_{i \neq i_0} r_i x_i$$

גם כאן - בהסתברות חצי $r_{i_0} = 0$ ואז:

$$\sum_{i} r_i y_i \neq \sum_{i} r_i x_i$$

ואליס תשיב נכונה גם כאן.
(הערה: כל הסכומים הם מודולו 2)
כיוון שמבצעים את הפרוצדורה פעמיים - ההסתברות לשגות בשתי הפעמים היא $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$, ולכן ההסתברות שענינו נכונה היא לפחות $\frac{3}{4}$, בסיבוכיות 3.

רנדומיות לא משותפת נניח כעת שמקור האקראיות אינו משותף, אלא יש מקור פרטי לכל אחד, והם בלתי תלויים, מה אפשר לעשות אז?
נתבונן אחרת בפרוטוקול הקודם - בוב יבנה טבלת אמת באורך 2^n, שבה כתוב, לכל מחרוזת באורך n - את חישוב הזוגיות של מחרוזת זו כפול y, אליס תשווה את שני הביטים שהיא מקבלת מול המקומות בטבלה שלה של כפל x בכל המחרוזות באורך n. נבחין כי התכונה שהשתמשנו בה - היה מובטח לנו שהטבלה של אליס והטבלה של בוב זהות במקרה ש $x = y$ והן שונות בדיוק במחצית מהתאים כאשר $x \neq y$.
האם נוכל לקוות לבנות טבלאות כנ"ל $C(x)$ ו $C(y)$ עבור אליס ובוב בהתאמה, שאכן יהיו שוות כאשר $x = y$ ושונות במחצית מהתאים כאשר $x \neq y$. האם נוכל לקוות שגודל הטבלאות יהיה קטן יחסית? כן, אפשר למשל לבנות משהו בסדר גודל של $10n$ ואז בוב יוכל להגריל 2 ביטים מהטבלה - לשלוח את האינדקסים שלהם ב $2 \cdot \log(10n)$,

ולקבל הסתברות שגיאה כמקודם, אחרי שאליס תשווה את הביטים שהתקבלו לשני ביטים במקומות המתאימים בטבלה שלה.

באופן יותר פורמלי:

פונקציה $C:\{0,1\}\to\Sigma^m$ נקראת קוד תיקון שגיאות עם מרחק δ, אם $\forall x,y\in\{0,1\}^n$ מתקיים:

$$x\neq y\Rightarrow Pr_{i\in\{1...m\}}\left[C\left(x\right)_i\neq C\left(y\right)_i\right]\geq\delta$$

כלומר - $C\left(x\right)$ ו $C\left(y\right)$ יהיו שונים ב $\delta\cdot m$ מהביטים.

טענה 1.2 קיים קוד תיקון שגיאות

$$RS:\{0,1\}^n\to\Sigma^m$$

עם $\delta\geq\frac{1}{2}$, כאשר $|\Sigma|\approx 10n$ וגם $|m|\approx 10n$. (ספציפית זהו קוד ריד־סולומון)

2 שיעור 2

מדוע זה פותר את הבעיה שלנו? בוב יקח את y, יחשב את $C\left(y\right)$, יבחר קואורדינטה i רנדומית במחרוזת $C\left(y\right)$ וישלח את $\left(C\left(y\right)_i,i\right)$, כלומר יאמר על איזו קוא' הוא מסתכל ומה ערכה. כיוון שערכי הקוא' הם $\{1,...,10n\}$ אז הוא ישלח $2\cdot\log\left(10n\right)$, וסך הכל זה מחייב $O\left(\log n\right)$ שימושים בערוץ. אליס בודקת האם $C\left(x\right)_i=C\left(y\right)_i$ (כאשר את האחד קיבלה, ואת השני היא יכולה לחשב). אם הם שווים אז היא משיבה ש $x=y$ ואחרת ש $x\neq y$. **הוכחה:** ידוע לנו שבין $5n$ לבין $10n$ קיים מספר ראשוני p (אפשר להוכיח את זה מהיות המקדם הבינומי $\binom{2n}{n}$ מספר שלם).

בהנתן $x\in\{0,1\}^n$ קיים פולינום $q\left(i\right)$ יחיד מדרגה לכל היותר $n-1$, המקיים $q\left(i\right)=x_i mod\, p$.

כעת נאריך את המילה שאנחנו רוצים לקודד (שארכה n) למילה באורך $n+p$, כאשר n הביטים הראשונים ישארו בעינם והביט ה $n+i$ יקבל את $q\left(n+i-1\right)$. וכך הביט ה 0 יקבל את ערך הפולינום עבור 0, הביט ה $n+5$ יקבל את $q\left(n+4\right)$ וכך הלאה.

מה אפשר לומר על המרחק של הקוד שהתקבל?

נניח שיש x,y, אז $C\left(x\right)_i=q_x\left(i\right)$ ו $C\left(y\right)_i=q_y\left(i\right)$.

אז הפולינום q_x-q_y הוא אפס בכל מקום שבו הקידודים שווים, אם $x\neq y$ אז זהו אינו פולינום האפס. מצד שני - זהו כן פולינום מדרגה לכל היותר $n-1$, ומשני אלו נובע שהפולינום מתאפס ב $n-1$ נקודות לכל היותר, ולכן כל שתי מילים שונות יעברו ע"י C למילים באורך $5n$ שהן מסכימות על חמישית מהקואורדינטות לכל היותר, ולכן המרחק של הקוד הוא $\frac{4}{5}$. ■

זהו קוד ריד סולופון.

תרגיל: לבנות קוד תיקון שגיאות עם $\delta>0$ קבוע, כך שאורך הקוד פולינומי ב n, והאלפבית נשאר בינארי.

2.0.3 רעש בערוץ

נניח שיש לנו קוד בוליאני כלשהו $C:\{0,1\}^n\to\{0,1\}^m$, שהוא קוד תיקון שגיאות עם מרחק $\delta\geq\frac{1}{4}$ ו $m\leq n^2$.

נניח לשם התרגיל שאת האינדקס i אפשר לשלוח בלי שגיאות, אבל שליחת $C\left(x\right)_i$ עלולה להיפגע. אז אפשר לשלוח פשוט את $C\left(x\right)_i$ מספר פעמים פרופורציוני להסתברות לשגיאה.

נניח כעת שהאינדקס עלול גם הוא להיפגע במהלך ההעברה, היות שיש בו $\log n$ ביטים, אז נשדר כל ביט בו $\log n$ פעמים. ואז במקום $\log n$ ביטים נשדר סך הכל $\log n\cdot\log n$ ביטים, ונקבל הסתברות הולכת וקטנה לשגיאה, כפונקציה של n.

אפשר במקום כך לבנות את הזוג $\left(C\left(x\right)_i,i\right)$ שארכו $\log n$, ואז לבנות את קוד תיקון השגיאות שלו ולשדר זאת, זה יצא $C\cdot\log n$ ביטים, ואם המרחק של קוד התיקון השגיאות שבנינו גדול מההסתברות לשגיאה בצורה משמעותית - נוכל להסיק מכך מה היתה המילה המקודדת לפני שהתרחשו בה שגיאות.

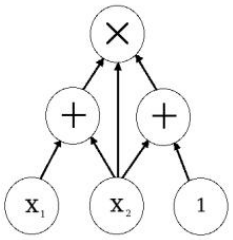
2.1 מעגל חישוב אלגברי

במעגל חישוב אלגברי יש:

- קבועים 1 ו -1

- משתנים (רגיסטרים) $x_1, ..., x_n$
- שערים לוגיים שמקשרים בין רגיסטרים לרגיסטרים, או קבועים, או בין תוצאות של שערים לוגיים.
- אחת התוצאות של השערים מוגדרת להיות תוצאת הפלט.

למשל (דיאגרמה ללא הקבוע -1):



ניתן גם להתבונן במעגל כפולינום, ולשאול למשל האם הוא זהותית שווה לאפס... נגיע לכך בהמשך.

הגדרה 2.1 גודל של מעגל C הוא מספר תוצאות הביניים במעגל.

.

הגדרה 2.2 בעיית ה $Identity$: בהנתן מעגל באורך n - בדוק האם הפולינום הפורמלי שהוא מחשב הוא פולינום האפס.

כלומר - הפולינום $x^2 - x^2$ הוא זהותית פולינום האפס, לעומת זאת, למרות שהפולינום $x^7 - x$ מתאפס לכל x מעל $\mathbb{Z}_7$, הוא אינו פולינום האפס.
אלגוריתם הסתברותי לפתרון: נחסום את דרגת הפולינום מלעיל באמצעות גודל המעגל, נבחר מספר כזה של קלטים, ונזין אותם לפולינום, אם נקבל אפס לכל קלט שכזה, נעריך שאכן זהו פולינום האפס זהותית.
לא ידוע אלגוריתם דטרמיניסטי.
(מדוע זו נקראת בעיית הזהות? כי בהנתן שני מעגלים, אם נרצה לבדוק האם זהותית מתקבל מהם אותו פולינום - אפשר את הפלט של אחד לכפול ב -1, לחבר אותו לפלט של השני, ולשאול האם מתקבל פולינום האפס).
בעצם למה זו בעיה קשה? כי למשל בפולינום (שהוא מעגל די קטן) $(x_1+1)(x_2+1)...(x_n+1)$ יש 2^n גורמים, לפתוח ולבדוק את המקדמים זה זמן אקספוננציאלי.
לבעיית חישוב הדטרמיננטה למשל יש מעגל ידוע בגודל פולינומי, נזכיר:

$$det(A) = \sum_{\sigma \in S_n} sign(\sigma) \cdot \prod_{i=1} a_{i,\sigma(i)}$$

לעומת זאת לבעיית הפרמננט, המוגדר:

$$per(A) = \sum_{\sigma \in S_n} \prod_{i=1} a_{i,\sigma(i)}$$

לא ידוע מעגל פולינומיאלי.
על מטריצת השכנויות של גרפים, הפרמננט נותן את מספר הזיווגים המושלמים בגרף הדו צדדי.

2.1.1 זיווג מושלם

בהנתן גרף G על הצמתים $V = \{1, ..., n\}$ נבנה מטריצה A :

$$(A)_{i,j} = \begin{cases} 0 & \{i,j\} \notin E(G) \\ x_{i,j} & j < i \wedge \{i,j\} \in E(G) \\ -x_{j,i} & i \leq j \wedge \{i,j\} \in E(G) \end{cases}$$

טענה 2.3 $det(A) \neq 0$ כפולינום במשתנים $x_{i,j}$ אם"ם קיים זיווג מושלם בגרף

הוכחה: בכיוון האחד - בהנתן גרף שיש בו זיווג מושלם, למשל גרף בן 8 קודקודים שיש את הזיווג $\{\{1,7\},\{2,5\},\{3,8\},\{4,6\}$, נבחין שבדטרמיננטה יתקבל הגורם $-x_{7,1}^2 \cdot x_{5,2}^2 \cdot x_{8,3}^2 \cdot x_{4,6}^2$ כפול סימן כלשהו של התמורה הרלבנטית. ואין שום גורם אחר שמאפס אותו, שכן גורם כנ"ל יוצא בדיוק מפרמוטציה אחת בדטרמיננטה.

בכיוון השני - נניח שאין זיווג. נבחין כי כל פרמוטציה מגדירה חלוקה של הגרף למעגלים מכוונים. נחפש גורם שאינו מתאפס ומתאים לתמורה π.

נבחין כי ב π חייב להיות מעגל באורך אי-זוגי, אחרת - אפשר לקבל חלוקה למעגלים זוגיים, לחלק את הגרף לגרף דו-צדדי ולקבל ממנו זיווג מושלם. את המעגל הזה - נהפוך בפרמוטציה, זה ישמר את הגורמים אך יהפוך את הסימן של התמורה. ■

3 שיעור 3

הגדרה 3.1 RP - אוסף בעיות ההכרעה L כך שקיים A אלגוריתם שרץ בזמן פולינומי, ועושה שימוש בביטים אקראיים, ולכל x מתקיים

$$x \in L \Rightarrow Pr[A(x) = Yes] = 1$$

$$x \notin L \Rightarrow Pr[A(x) = No] \geq \frac{2}{3}$$

(שני שליש הוא כאן קבוע שרירותי).

קל לראות שע"י חזרות אפשר לשפר את $\frac{2}{3}$ למשהו שהוא קרוב ל1 במידה שהיא אקספוננציאלית ב n, למשל $1-10^{-n}$. יהי n.

נתבונן בהסתברות על פני הביטים האקראיים:

$$Pr[\exists x, |x| = n,\ A(x) \neq L(x)]$$

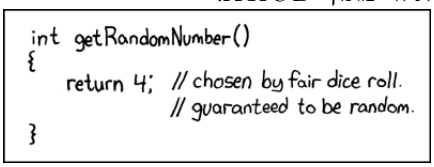
מכיוון שנוכל לבחור אלג' כך שההסתברות עבור x ספציפי לשגיאה אינה עולה על 10^{-n} (לפי ההערה הקודמת), אז ההסתברות שקיים x כנ"ל היא 2^n (מספר הקלטים האפשריים) כפול ההסתברות לעיל. מכאן שההסתברות הנ"ל חסומה ע"י

$$10^{-n}2^n = 5^{-n}$$

נובע שכמעט כל מחרוזת רנדומית שניקח תעבוד עבור כל הקלטים בצורה מושלמת.

שאלה: האם יש אלג' פולינומי G כך שבהנתן n מייצר מחרוזת r כך ש A_rעובד עבור כל הקלטים מגודל n? אם כן - נריץ את G, נקבל מחרוזת, נריץ את A עם המחרוזת שהתקבלה ונקבל תשובה נכונה בהסתברות 1. ולכן $RP \subseteq P$.

איור מתוך $XKCD$:



מסקנה 3.2 אם היה לנו מחולל G כנ"ל היינו מוכיחים $RP \subseteq P$.

הגדרה 3.3 נגיד כי שפה L היא ב $P/Poly$ אם קיים אלג' דטרמיניסטי פולינומיאלי A, המקבל שני קלטים x,r כך שמתקיים:
קיים פולינום $p(n)$ כך שלכל n מתקיים $|r| \leq p(n)$ ולכל n יש $r = r(n)$ מתאים, כך שמתקיים:

$$\forall |x| = n,\ A(r,x) = L(x)$$

בעצם יש ב $P/Poly$ שפה בלתי כריעה. אם ניקח את בעיית העצירה ונקודד את המכונות בקידוד אונרי - אז יש שפה אחת מכל גודל ונצטרך מחרוזת באורך ביט 1 שתאמר לנו האם המחרוזת 1^m מקודדת מכונה שעוצרת או לא, ולכן $P/1$ מכילה שפה בלתי כריעה...

מגדירים את הדרגה של פולינום האפס להיות $-\infty$, ודרגה של $x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot ... \cdot x_n^{\alpha_n}$ היא $\sum \alpha_i$. באופן דומה דרגה של פולינום מרובה משתנים היא דרגת המחובר בעל הדרגה המקסימלית.

משפט 3.4 הלמה של ציפל שוורץ. יהא $\mathbb{F}$ שדה סופי ויהי $p \in \mathbb{F}[x_1,...,x_n]$ פולינום כך ש $deg(p) \leq d$, שאינו פולינום האפס.
אזי

$$Pr_{x_1,...,x_n \in \mathbb{F}}[p(x_1,...,x_n) = 0] \leq \frac{d}{|\mathbb{F}|}$$

הוכחה: באינדוקציה על מספר המשתנים. מקרה הבסיס מהתאפסות פולינום במשתנה יחיד וכו'.
יהי p פולינום ב n משתנים מדרגה d לכל היותר. נרשום את p כך:

$$p(x_1,...,x_n) = q_0(x_1,...,x_{n-1}) + x_n \cdot q_1(x_1,...,x_{n-1}) + x_n^2 \cdot q_2(x_1,...,x_{n-1}) + ... + x_n^{d'} q_{d'}(...,x_{n-1})$$

וברור ש $d' \leq d$.
אם $d' = 0$ אז x_n לא מופיע, והדרוש מתקיים מהנחת האינדוקציה.
נניח $d' \neq 0$ ו $q_{d'} \neq 0$ ונסמן את הדרגה של $q_{d'}$ באות l. מכאן שהפולינום $x_n^{d'} q_{d'}$ הוא מדרגה $l + d'$, והפולינום כולו הוא מדרגה d לכל היותר מתקיים:

$$l + d' \leq d$$

כל q_k יתאפס בהסתברות שהיא $\frac{deg(q_k)}{|\mathbb{F}|}$, ולכן q_l יתאפס בהסתברות $\frac{l}{|\mathbb{F}|}$. כעת אם נקבע את $x_1,...,x_{n-1}$ ונשאל מה ההסתברות להתאפסות p, על פני x_n - נקבל שזהו פולינום מדרגה d', ולכן ההסתברות להתאפסות היא $\frac{d'}{|\mathbb{F}|}$ לכל היותר. כלומר מנוסחת ההסתברות השלמה:

$$\begin{aligned} Pr_{x_1,...,x_n}[p(x_1,...,x_n) \neq 0] &\geq Pr_{x_1,...,x_{n-1}}[q_{d'}(x_1,...,x_{n-1}) \neq 0] \cdot Pr_{x_n}[P(x_1,...,x_n) \neq 0 \,|\, q_{d'}(x_1,...,x_{n-1}) \neq 0] \\ &\geq \left(1 - \frac{l}{|\mathbb{F}|}\right) \cdot \left(1 - \frac{d'}{|\mathbb{F}|}\right) \\ &\geq 1 - \frac{d' + l}{|\mathbb{F}|} \\ &\geq 1 - \frac{d}{|\mathbb{F}|} \end{aligned}$$

כנדרש. ■

.

משפט 3.5 בעית הזהות עבור מעגלים אלגבריים היא ב RP

הוכחה: מספיק להכריע האם מעגל אלגברי מחשב את פולינום האפס.
דרגת הפולינום חסומה מלעיל ע"י 2^n (n גודל המעגל).
נבחר שדה גדול מהדרגה לעיל, נבחר אקראית מספרים בתחום וניצב אותם בפולינום, כיוון שידוע לנו חסם על ההסתברות, אם נקבל מספיק פעמים 0 נוכל להכריז שהפולינום הוא פולינום האפס.
יש כמה חורים בהוכחה - כמו למשל "איך מוצאים שדה כל כך גדול" ואחרי שמוצאים עלולים להיות קשיים אם המציין של השדה מחלק את מקדמי הפולינום (שאז נקבל 0 תמיד גם אם הפול' אינו פול' האפס). ■

3.1 מעט רקע על קודים בינארי לתיקון שגיאות

קוד לתיקון שגיאות היא קבוצה $C \in \{0,1\}^n$. לאיברי C קוראים מילות הקוד, האורך של Cהוא n.
רוצים ניצול יעיל של ערוץ התקשורת ויכולת טובה של תיקון שגיאות.
נשדר רק מילות קוד חוקיות, ואם התרחש רעש בדרך, מי שמקבל את המידע, ישער על סמך היכרות עם C את מילת הקוד הקרובה ביותר למילת קוד חוקית (ביחס למרחק המינג).
אם המרחקים ההדדיים בין מילות הקוד גדולים, אז שיטת הפיענוח על סמך קירבה - תיתן תוצאות טובות. קונקרטית, אם המרחק המינימלי בין שתי מילים הוא d, אז לכל מילה ששובשה בפחות מ $\frac{d}{2}$ מהקואורדינטות שלה - נקבל פיענוח נכון.
ספציפית - מגדירים את המרחק של C להיות המרחק המינימלי בין שתי מילים בקוד.
[מכאן אני קצת מזניח את הנושא, מי שלא מתמצא מוזמן להסתכל בסיכומים ממבוא לתורת האינפורמציה, או בויקיפדיה על קודי המינג וקודים ליניאריים]
R מוגדר להיות הקצב של הקוד $\frac{1}{n}\log_2|C|$
נגדיר את המרחק המנורמל $d(C)$ להיות המרחק המינימלי לחלק ל n (באשר n הוא מספר האותיות במילת קוד חוקית).
רוצים ש R ו $d(C)$ יהיו גדולים ככל האפשר, ושהפיענוח יהי יעיל.
נזכיר ש C הוא קוד לינארי כאשר C היא לא סתם קבוצה אלא תת מרחב לינארי של Σ^n.
מטריצת בדיקת הזוגיות מגודל $k \times n$ היא מטריצה שהגרעין שלה הוא בדיוק C, ואז מתקיים $R = 1 - \frac{k}{n}$.
המרחק (הלא מנורמל) של C הוא האורך הקטן ביותר של תלות ליניארית בין עמודות A (למשל באלפבית בינארי - מספר העמודות המינימלי שאפשר לקבל מחיבורן 0).

4 שיעור 4

ראינו את קוד $Hamming$, שהוא קוד ליניארי $[2^t-1, 2^t-t-1, 3]_q$

הגדרה 4.1 אומרים שקוד C הוא מושלם אם יש פרמטר γ כך שכל הכדורים ברדיוס $\frac{d}{2}$ מסביב למילות קוד חוקיות, הם זרים, ואיחודם הוא $\{0,1\}^n$.

נשים לב שקוד המינג הוא מושלם.
זרות הכדורים: היות שכדורים ברדיוס אחד סביב מילות קוד חוקיות הם זרים.
מילוי המרחב: מהו מספר מילות הקוד? זהו $2^{n-t} = \frac{2^n}{2^t} = \frac{2^n}{n+1}$
בכדור ברדיוס יחידה יש $n+1$ מילים בדיוק, ולכן סכום הנפחים של הכדורים:

$$(n+1)\frac{2^n}{n+1} = 2^n$$

והם אכן ממלאים את המרחב.
ידוע שמלבד קודי המינג (ומעל א"ב כלשהו) יש רק עוד שני קודים מסויימים (קודי $Golay$) מושלמים ואין עוד מלבדם.

4.1 מציאת קוד

בעיה כללית: בהנתן n, d למצוא את $A(n,d)$ הגודל המירבי של קוד מאורך n ומרחק לפחות d.
גירסה אסימפטוטית (למקרה ש $d \sim n$)

$$R(\delta) = \limsup_{n\to\infty} \{R(C) \,|\, C \subseteq \{0,1\}^n,\ d(C) \geq \delta \cdot n\}$$

חסם הכדורים - ניקח את מספר המילים במרחב, ונחלק אותם במספר המילים המוכל בכדור ברדיוס d, ונקבל חסם עליון למספר המילים בקוד (כיוון שכל כדור מכיל מילת קוד יחידה):

$$|C|\left(\sum_{i\leq\frac{d}{2}}\binom{n}{i}\right)\leq 2^n$$

נזכיר כי:

$$\binom{n}{\alpha n}\approx 2^{n(H(p)+o(1))}$$

ואז נחליף את המחובר הכי משמעותי בסכום, שהוא $\binom{n}{\frac{d}{2}}$, ב $2^{n\cdot H\left(\frac{\delta}{2}\right)}$, ומכיוון שגדלי המקדמים הבינומיים הם אקספוננציאליים, אז הוא בעצם קובע את ההתנהגות של הסכום כולו, ולכן אפשר להתעלם מהמחוברים האחרים ולקבל:

$$|C|\leq 2^n\cdot 2^{-nH\left(\frac{\delta}{2}\right)}$$

ולכן:

$$R\leq 1-H\left(\frac{\delta}{2}\right)$$

וחסם הכדורים הוא :

$$R\left(\delta\right)\leq 1-H\left(\frac{\delta}{2}\right)$$

4.2 חסם אליאס

נתחיל מכך שאם $\delta\geq\frac{1}{2}$, $R\left(\delta\right)=0$. במילים אחרות - אם דורשים שכל שתי מילות קוד יבדלו זו מזו במחצית מארכן - אזי לא נוכל שמספר מילות הקוד יהיה אקספוננציאלי. בפרט - קוד הדמר שראינו, ונותן מספר ליניארי של מילים, הוא אופטימלי במובן זה.

טענה 4.2 אין קוד מאורך n בגודל מעריכי ב n שבו המרחק הממוצע בין מילות קוד הוא $\frac{n}{2}$.

הוכחה: נתבונן במטריצה $|C|\times n$ ששורותיה הן מילות הקוד החוקיות, ונחשב את המרחק הממוצע בין שתי מילות קוד:

$$\frac{1}{\binom{|C|}{2}}\sum_{x\neq y\in C}d\left(x,y\right)$$

ובעצם לפי הגדרת מרחק המינג זה גם:

$$\frac{1}{\binom{|C|}{2}}\sum_{x\neq y\in C}\#\left\{i|x_i\neq y_i\right\}$$

ובהיפוך סדר הסכימה:

$$\frac{1}{\binom{|C|}{2}}\sum_{i}\#\left\{(x,y)\ as\,before|x_i\neq y_i\right\}$$

אם נסמן ב N_i וב Z_i את מספר ה 1 וה 0 בהעמודה ה i בהתאמה, אזי הסכום לעיל הוא בדיוק:

$$\frac{1}{\binom{|C|}{2}}\sum_i N_i Z_i$$

ומכיוון ש $N_i+Z_i=|C|$ אז מכפלתם מקסימלית כאשר $N_i=Z_i=\frac{|C|}{2}$ ולכן הסכום לעיל קטן מ:

$$\frac{1}{\binom{|C|}{2}}\sum_i \left(\frac{|C|}{2}\right)^2$$

המקדם הבינומי הוא בערך ריבועי ב $|C|$ולכן זה בקירוב $\frac{n}{2}$. ■

משפט 4.3 חסם אליאס $(Elias)$:

$$R(\delta)\leq 1-H\left(\frac{1-\sqrt{1-2\delta}}{2}\right)$$

נרחיב את ההוכחה לעיל לחסם של אליאס:
ננית שבמטריצה הזו, שיעור ה 1־ים הוא p כך ש $0<p\leq\frac{1}{2}$.
תרגיל פשוט מראה שעל מנת למקסם את $\sum N_i Z_i$ עדיף ש $\forall i\ N_i\simeq p\cdot|C|$
ולכן, אם ידוע לנו שהמשקל הממוצע של מילת קוד הוא $p\cdot n$, אז בהוכחה הקודמת נוכל להחליף את N_iZ_i ב $2p(1-p)n$ ואז נקבל שהמרחק הממוצע ב C לא יעלה על $|C|^2 p(1-p)$.

לסיכום: אם C קוד מאורך n כך שהמשקל הממוצע של מילת קוד ב C הוא $p\cdot n$, ואם הגודל של $|C|$ מעריכי, אזי המרחק הממוצע בין מילות הקוד הוא לכל היותר $2p(1-p)n$ כנ"ל. בעצם לא מוכרחים ש C יהיה מעריכי בגדלו, ודי שיהיה ריבועי. **הוכחה:** (חסם אליאס)

נחתוך את C עם ספירת המינג אקראית $S=S\left(\underbrace{x}_{central\,point},\underbrace{r}_{radius}\right)$ ברדיוס מתאים.

מהי התוחלת של גודל חיתוך C ו S?

$$\mathbb{E}_S[C\cap S]=\frac{|C|\cdot|S|}{2^n}$$

מדוע זה נכון? נבחר את מרכז הספירה באקראי (את הרדיוס קבענו מראש) ואז התוחלת תהיה:

$$\frac{1}{2^n}\sum_x|C\cap S(x,r)|=\frac{1}{2^n}\sum_{x,y}1_{y\in C\wedge d(x,y)=r}$$

בהיפוך סדר הסכימה:

$$\begin{aligned}\frac{1}{2^n}\sum_{y\in C}\#\{x|d(x,y)=r\}=&\frac{1}{2^n}\sum_{y\in C}\binom{n}{r}\\=&\frac{|C|\cdot\binom{n}{r}}{2^n}\\=&\frac{|C|\cdot|S|}{2^n}\end{aligned}$$

מכאן נובע שיש ודאי ספירה ספציפית שנותנת לפחות את המשקל הממוצע.

כלומר תהי S ספירה ברדיוס r כך ש

$$|S \cap C| \geq \frac{|C| \cdot \binom{n}{r}}{2^n}$$

נתבונן רק במילות הקוד החוקיות שעל הספירה.
ניקח את כל מילות הקוד בחיתוך עם הספירה, נפעיל עליהן xor עם מרכז הספירה x, זה בעצם שקול להזזת הקבוצה הזו כך שמרכז המעגל ימצא בראשית הצירים. עכשיו כל מילות הקוד (ב $S \cap C$ לאחר ההזזה) משקלן בדיוק r.
כדי לקבל את חסם אליאס, צריך לבחור r אופטימלי (אם בוחרים $r = \rho n$ אז $\delta = 2\rho(1-\rho)$ ומקבלים מכך $\rho = \frac{1-\sqrt{1-2\delta}}{2}$. ■

4.3 חסם גילברט ורשמוב

נראה כעת חסם תחתון, כלומר חסם המראה שיש קודים טובים.
והוא:

$$R(\delta) \geq 1 - H(\delta)$$

השיקול החמדני: משמיטים כנגד כל מילת קוד כדור המינג ברדיוס $\delta \cdot n$ מסביבה, נפחו הוא $\sum_{i \leq \delta n} \binom{n}{\delta}$, זה כאמור אסימפטוטית $2^{n(H(\delta)+o(1))}$ ולכן התהליך נמשך:

$$\frac{2^n}{2^{nH(\delta)}} = 2^{n(1-H(\delta))}$$

איטרציות, ולכן זה מספר מילות הקוד, וחסם תחתון לקצב מתקבל מהוצאת לוג וחלוקה ב n כמקודם.
בעצם - יש אפילו קוד ליניארי בקצב $1 - H(\delta)$ ומרחק מנורמל שהוא לפחות δ.
מספיק להציג מטריצת בדיקת זוגיות מתאימה, כלומר מטריצה בינארית $k \times n$ שאין בה קבוצה קטנה של עמודות תלויות - כלומר שהתלות הקצרה ביותר גדולה מ δn.
את k נקבע בהמשך, נזכיר רק שבכל מקרה מתקבל $R = 1 - \frac{k}{n}$.
ברור שאסור לקחת את עמודת ה 0, ולכן יש $2^k - 1$ מועמדים לעמודות.
אחרי שבחרנו עמודה, נוכל לבחור כל עמודה אחרת שנרצה, כלומר יש $2^k - 2$ אפשרויות.
אחרי שיש שתי עמודות, נוכל לבחור כל עמודה שאינה צירוף ליניארי של שתיהן, שהיינו יש $2^k - 4$ אפשרויות.
באופן כללי - עבור המקום ה $n+1$ - אסור לבחור אף אחת מהעמודות שכבר נבחרו, אסור לבחור אף עמודה שהיא סכום של שתי עמודות, אסור לבחור אף עמודה שהיא סכום של שלוש עמודות, וכן הלאה עד סכומים של $d-1$ עמודות. ולכן, אם הקשר בין n ו k מאפשר בחירה שכזו, כלומר:

$$0 < 2^k - \sum_{i<d} \binom{n}{i}$$

אז אפשר באמצעות האלגוריתם הנ"ל לבנות את מטריצת בדיקת הזוגיות המבוקשת.
כמקודם, המקדם הבינומי המשמעותי הוא האחרון ולכן נדרוש

$$0 < 2^k - 2^{nH(\delta)} \Rightarrow 2^{nH(\delta)} < 2^k$$
$$\Rightarrow k < nH(\delta)$$

ובמידה שזה מתקיים - נוכל לבנות קוד ליניארי כנ"ל.

4.3.1 שתי הערות על גיאומטריה במימדים גבוהים

כדורים בשני מימדים - כדור היחידה תופס בערך 0.78 מקובית היחידה.
בשלושה מימדים - נפח הקוביה בערך 8 ונפח הכדור $\frac{4\pi}{3}$ ואז הכדור תופס 0.52 מהקוביה.
ככל שהמימד עולה - היחס הזה דועך מעריכית, שכן:

$$V_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} = \begin{cases} \frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!} & even\, n \\ 2 \cdot \frac{(2\pi)^{\lfloor \frac{n}{2} \rfloor}}{n!!} & odd\, n \end{cases}$$

אם ניקח חגורה ברוחב ε סביב קו המשווה - ככל שנעלה במימד, החגורה תהיה כמעט כל הכדור, כלומר החלק מהספירה שאינו החגורה - גדלו דועך מעריכית.

סריגים (lattices) ניקח קבוצה במישור, הנפרשת ע"י השלמים ושני וקטורים, כלומר:

$$l = \{n_1 x_1 + n_2 x_2 | n_1, n_2 \in \mathbb{Z}\}$$

בעיה: בהנתן סריג n מימדי l, רוצים לדעת מהו הוקטור הקצר ביותר בו. היות ש 0 בסריג, אז זו שאלה שקולה למרחק הקצר, ונראה שיש קשר בין זה לבין שאלת המרחק של קודים...

4.4 קודי הדמר

4.4.1 אי שויון הדמר

תהי $A_{n\times n}$ ויהיו $a_1, ..., a_n$ וקטורי השורות בה.
אז:

$$|detA| \leq \prod_{i=1}^{n} \|a_i\|_2$$

הדטרמיננטה היא נפח המקבילון הנפרש ע"י הוקטורים, ומקבלים נפח מקסימלי כאשר הוקטורים ניצבים, וזה בדיוק הביטוי הנתון באגף ימין.

4.4.2 מטריצות הדמר

מתבקש להבין את הדוגמאות שבהן מתקיים שוויון. קל לראות שזה קורה אם"ם a_i ניצבים זל"ז.
האם זה יכול לקרות כאשר a_iהם וקטורי ± 1?

הגדרה 4.4 מטריצה $n \times n$ שאיבריה ± 1 ושורותיה ניצבות זל"ז נקראת מטריצת הדמר.

באופן שקול A מטריצת הדמר אם כל שתי שורות שלה מסכימות בדיוק על מחצית מאיבריהן.
וכן $A \cdot A^T = n \cdot I$ (כיוון שמכפלת מטריצה במשוחלפת שלה נותנת במקום ה i,j את המכפלה הפנימית של העמודה ה i ב j)
לא קשה להראות שאם $n \geq 4$ויש מטריצת הדמר $n \times n$ אז בהכרח $4|n$.
השערה פתוחה: לכל n המתחלק ב 4 יש מטריצת הדמר $n \times n$.

בניית סילבסטר בניית סילבסטר מספקת טכניקה פשוטה לבנות מטריצות הדמר מסדר 2^n, מתוך ההבחנה שאם H מטריצת הדמר אז גם $\begin{pmatrix} H & H \\ H & -H \end{pmatrix}$.
באופן יותר טכני - אם $n = 2^t$ אז ניתן לכל שורה מספר ע"י וקטור בינארי באורך t, ואז עבור השורה המאונדקסת ע"י x והטור y מגדירים את התא $a_{x,y} = (-1)^{\langle x,y \rangle}$.

בניית Paley יהי q חזקה אי־זוגית של מספר ראשוני $q=p^{\alpha}$, אז יש שדה $\mathbb{F}=GF(q)$. ועליו נגדיר את הכרקטר הריבועי χ. זוהי העתקה $\chi:\mathbb{F}\to\{0,-1,1\}$ כך:

$$\chi(x)=\begin{cases}0 & x=0\\ +1 & x\,is\,a\,square,\ \exists t\in\mathbb{F}:t^2=x\\ -1 & x\,is\,not\,a\,square\end{cases}$$

ניקח מטריצה ששורותיה ועמודותיה מאונדקסות ע"י איברי השדה ובתא ה x,y נרשום את $\chi(x-y)$. אזי על האלכסון רשומים אפסים, אבל בשאר המקומות רשום 1 ו -1, וכל השורות ניצבות זל"ז.
כעת אם $q\equiv 3\,(mod\,4)$, רושמים באלכסון 1 במקום 0, ומוסיפים שורה ועמודה של 1ים, אז מתקבלת מטריצת הדמר.

4.5 .

פולינומים נותנים כלי יעיל בתיאור ובבניה של קודים לתיקון שגיאות. בהנתן P פולינום המוגדר ע"י מקדמים או ערכים בנקודות (במספר מתאים) אזי אפשר לקודד את הפולינום ע"י חישובו בכל הערכים שהם המילים, כלומר $(P(\vec{y}))_{\vec{y}}$. מקבלים כך קידוד יתר, כלומר יותר מידע ממה שדרוש, ובדיוק בו נוכל להשתמש כדי לשחזר את הפולינום במידה שאירעה שגיאה.
ידוע לנו ש $d+1$ נק' במישור קובעות ביחידות פולינום אינטרפולציה אחד ויחיד ממעלה $\geq d$, ואם נקבל קבוצה גדולה מאד של נקודות, אז כל $d+1$ נקודות יתנו לנו מועמד לפולינום אינטרפולציה, נרצה למצוא דרכים לשחזר את הפולינום המקורי.

5 שיעור 5

(סיכם והקליד: אלון גונן)
הראינו שאין קודים גדולים עם מרחק מנורמל $>\frac{1}{2}$. נראה כרגע ונוכיח (באופן מדוקדק הפעם) את חסם ג'ונסון.

5.1 חסם ג'ונסון

הערה 5.1 נחזור על הגדרה: עבור $x\in\{0,1\}^n$, $B(x,(\frac{1}{2}-\epsilon)\cdot m)$ זה הכדור סביב x עם רדיוס (לא מנורמל) $(\frac{1}{2}-\epsilon)\cdot m$.

משפט 5.2 יהי קוד $C\subseteq\{0,1\}^n$ במרחק מנורמל $\leq(\frac{1}{2}-\epsilon)$ ויהי $\delta\geq\sqrt{\epsilon}$. אז לכל $x\in\{0,1\}^n$ מספר מילות הקוד ב־$B(x,(\frac{1}{2}-\delta)\cdot m)$ אינו עולה על $\frac{1}{2\cdot\delta^2}$.

הוכחה: תהינה $y_1,...,y_l\in C$ כך שמרחקן מ־x $\geq(\frac{1}{2}-\delta)\cdot m$. נסמן $z_i=x\oplus y_i$. נשנה מעט את האינטרפטציה של xor. xor עבור שני ביטים זהים יתן 1 ועבור ביטים שונים יתן -1. נשים לב כי לאור הנחותינו מתקיים

$$\forall i\ \ <z_i,1>\geq(\frac{1}{2}+\delta)\cdot m-(\frac{1}{2}-\delta)\cdot m=2\cdot\delta\cdot m$$

בגלל הנחותינו על המרחק מתקיים

$$\forall i\neq j\qquad <z_i,z_j>=<y_i,y_j>\leq(\frac{1}{2}+\epsilon)\cdot m-(\frac{1}{2}-\epsilon)\cdot m=2\cdot\epsilon\cdot m$$

נגדיר $w=\Sigma_{i=1}^{l}z_i$. מתקיים

$$||w||_2=\Sigma||z_i||+2\cdot\Sigma_{i\neq j}<z_i,z_j>\leq l\cdot m+2\cdot l^2\cdot\epsilon\cdot m$$

מצד שני, לפי קושי־שוורץ

$$||w||_2^2 \geq \frac{1}{m} \cdot < w, 1 >^2 = \frac{1}{m} \cdot (\Sigma < z_i, 1 >)^2 \geq \frac{1}{m} \cdot (l \cdot 2 \cdot \delta \cdot m)^2 = 4 \cdot \delta^2 \cdot l^2 \cdot m \geq 4 \cdot \epsilon \cdot l^2 \cdot m$$

לסיכום, נקבל

$$l \cdot m + 2 \cdot l^2 \cdot \epsilon m \geq 4 \cdot l^2 \cdot \epsilon \cdot m$$

ונקבל $l \leq \frac{1}{2 \cdot \delta^2}$ ■

5.2 שירשור קודים

ראינו את קודי RS (ריד־סלומון). הראינו לגביו כי המרחק המנורמל שלו הוא $1 - \frac{n}{m}$. זהו קוד ליניארי. ניתן להראות שבסיס עבור RS ניתן ע"י עמודות המטריצה

$$\begin{pmatrix} 1 & z_1 & . & . & z_1^{n-1} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & z_m & . & . & z_{m^{n-1}} \end{pmatrix}$$

משפחה אחרת של קודים זה קודי RM. שם היה לנו שדה F, $d, l \in \mathbb{N}$ כאשר $d < |F|$. קוד RM ניתן לראות כאוסף כל הפולינומים ממעלה כוללת $\geq d$ ב־l משתנים ולכן זה קוד ליניארי.
נשאל כמה פולינומים ממעלה כוללת $\geq d$ ב־l משתנים יש. המספר הוא $|F|^{number \text{of monoms}}$ (לכל מונום עלינו לבחור מקדם). מספר המונומים הוא לכל היותר $\begin{pmatrix} l+d \\ l \end{pmatrix}$. בסך הכל נקבל כי מספר הפולינומים הוא לכל היותר $|F|^{\begin{pmatrix} l+d \\ l \end{pmatrix}}$.
כאשר מתאימים כל פולינום שכזה לכל ערכיו (יש לנו $|F|^{|F|^l}$ כאלה) מתקבלת העתקה

$$|F|^{\begin{pmatrix} l+d \\ l \end{pmatrix}} \to |F|^{|F|^l}$$

זה בעצם קוד RM. נבחין שכאשר $l = 1$, $F = GF(2)$ מתקבל קוד RS.
כפי שראינו המרחק המנורמל של קוד RM הוא $\leq 1 - \frac{d}{|F|}$.
קוד הדמר (WH) הוא הטוב ביותר האפשרי בתחומו מבחינת המרחק אבל גרוע מבחינת הקצב. לקוד RS יתרון במרחק וקצב אך חסרונו בכל שאינו בינארי. נראה איך לצרפם ביחד ע"י פעולת שירשור.
נגדיר את השירשור באמצעות ההעתקה הבאה:
.

$$x \to WH(RS(x)_1)...WH(RS(x)_m)$$

בדומה ניתן לשרשר יחד קודים שונים.

המרחק של RS הוא $\delta_1 = 1 - \frac{n}{m} \leq$.
המרחק של WH הוא $\delta_2 = \frac{1}{2}$.
ניווכח כי לקוד המשורשר שהגדרנו מרחק $\delta_1 \cdot \delta_2$. לשם כך נעבור לנוטצייה של העתקות. יש לנו שתי העתקות:

$$RS : F^n \to F^m$$

$$WH : \{0,1\}^{log|F|} \to \{0,1\}^{|F|}$$

הקוד המשורשר $WH \circ RS$ אם כך יוגדר

$$WH \circ RS : \{0,1\}^{n \cdot log|F|} \to \{0,1\}^{m \cdot |F|}$$

יהיו $x, y \in \{0,1\}^{n \cdot log|F|}$ ונביט בתמונות שלהן עבור ההעתקה $WH \circ RS$. $RS(x), RS(y)$ שונים ב־$\delta_1 \cdot m$ סימבולים. אם $RS(x)_j \neq RS(y)_j$ אזי WH שלהם שונה בשיעור δ_2 מהמקומות. סה"כ המחרוזות שונות לפחות בשיעור $\delta_1 \cdot \delta_2$.
נשאל כעת מה יצא לנו מהשירשור הזה?
כפי שכבר ציינו לכל $n \in \mathbb{N}$ קיים מספר ראשוני בין n ל־$2n$. לכן יש הרבה שדות סופיים. כשמצרפים אותם ע"י שירשור כנ"ל אזי ניתן לקבל קודים עם תהליך קידוד פולינומי $\{0,1\}^n \to \{0,1\}^{20n}$ במרחק מנורמל ≤ 0.4.

5.3 פיענוח קודים

נראה אלגוריתם לפיענוח יעיל של קודי RS.

5.3.1 אלגוריתם $Berlekamp\,Walch$ לפיענוח קודי RS

כידוע, פולינומים ממעלה d מעל שדה נקבע ביחידות ע"י $d+1$ נקודות על העקום המתאים. השאלה היא אם כן בעלת האופי הבא: נתונים לנו ערכי p בנקודות שונות עם שיעור קבוע של טעויות. האם עדיין ניתן לפתור את בעיית האינרפולציה (לשחזר את p).

משפט 5.3 יש אלג' פולינומי כדיקלמן: הוא מקבל כקלט רשימה $(a_1, b_1), ..., (a_m, b_m)$ (נשים לב כי המקרה המעניין הוא כאשר $m > d$) כאשר לכל $i \in [m]$, a_i, b_i איברים בשדה סופי כך שיש פולינום G ממעלה $\leq d$ המקיים $G(a_i) = b_i$ לפחות עבור $\frac{m+d}{2}$ ערכי i. האלגוריתם מוצא את G בזמן פולינומיאלי.

הוכחה: נתחיל מ"חימום":
אם מספר האינדקסים שבשבילם $G(a_i) = b_i$ הוא $m \cdot (1 - \frac{1}{2d+2})$. אז השיטה הפשוטה הבאה תעבוד. נבחר $d+1$ ערכים אקראיים של i ונחשב את פולינום האינטרפולציה המתאים. קל לראות כי תחת ההנחה ה"ל נמצא את הפולינום בהסתברות $\leq \frac{1}{\sqrt{e}}$.
כעת נחזור להוכחה. נראה גירסה מוחלשת (העיקרון של ההוכחה המלאה דומה). נניח לשם הפשטות כי $m = 4d$ והשיוון $g(a_i) = b_i$ מתקיים לפחות ל־$t = 3d$ ערכי i.
נראה שיש פולינום C ממעלה $\geq 2d$ ופולינום E כך ש־E הוא לא זהותית פולינום האפס. ממעלה $\geq d$ כך שמתקיים

$$\forall i \quad C(a_i) = b_i \cdot E(a_i)$$

כלומר, נראה
א. שיש פולינומים כאלה.

ב. נראה שהפולינום E מחלק את C ואז פשוט נגדיר $G = \frac{C}{E}$ (במקומות שיש טעויות E יתאפס ואז לא נוכל לתקן בצורה הזו).

הוכחת א': (אם נוותר על התנאי ש־E לא זהותית 0 אז אם נגדיר גם את C להיות פולינום האפס נקיים את הדרישה של א'). יש קבוצה מגודל $d \geq$ (בלתי ידועה לנו) עבורה $G(a_i) \neq b_i$. נדרוש מהפולינום E להתאפס ב־d הנקודות האלה. ניתן לספק את הדרישה הזו באופן לא טריוויאלי (פולינום E לא זהותית אפס). בבחירה כזו של E^* מתאים אנו יכולים לבחור $C = G \cdot E$ ואז מתקיימים כל תנאי א'. יוצא אם כן ש־

$$C(x) = G(x) \cdot E(x)$$

מתקיים לפחות ל־$3d$ ערכים שונים של x ולכן $G = \frac{C}{E}$.

השיוויון הזה מתקיים בכל נקודה שבה $g(a_i) = b_i$. ל־x כזה מתקיים $C(a_i) = b_i \cdot E(a_i) = G(a_i)$ ■

5.4 פיענוח מקומי *local decoding*

נתון קוד $C : \{0,1\}^a \to \{0,1\}^b$. מקבלים $y \in \{0,1\}^b$ ומחפשים $x \in \{0,1\}^a$ כך ש־$C(x) \sim y$. נוסיף עוד אספקט: מלבד y נקבל אינדקס j ורוצים לדעת את x_j בהסתברות $\geq \frac{2}{3}$.

נראה שתי דוגמאות:

5.4.1 פיענוח מקומי לקוד הדמר

נזכיר כי קוד הדמר ממפה $x \in \{0,1\}^n$ ל־$(< x, y_1 > ...., < x, y_{2^n} >)$ כאשר ערכי y_i עוברים כל הווקטורים ב־$\{0,1\}^n$. אנחנו רוצים כאמור למצוא את x_j. אחד ה־y הוא e_j. ואז $x_j =< x, e_j >$. נראה איך ניתן לפענח מקומית את WH כך שלכל $p < \frac{1}{4}$ לגלות את x_j בהסתברות $\frac{1}{3} < 1 - 2p$. קיימים y_1, y_2 כך ש־$y_1 \oplus y_2 = e_j$. אם הקורדינטות y_1, y_2 ב־$WH(x)$ נכונות אז אפשר לחשב:

$$x_j =< x, e_j >=< x, y_1 + y_2 >=< x, y_1 > + < x, y_2 >$$

נבחר את y_1 באקראי ואז נגדיר $y_2 = y_1 \oplus e_j$ (ונקבל y_1, y_2 כנדרש). ניתן לראות שאם נמשיך כך סיכויי ההצלחה הם $\leq 1 - 2p$.

5.4.2 פיענוח מקומי לקוד RM

עכשיו נחשוב על קוד RM כהרחבה של פולינום במשתנים מרובים p מתת קבוצה של F^l לקב' גדולה. ידועים לנו (בדרגשת שיבוש ρ) ערכיו של פולינום ממעלה כוללת $d \geq$ ב־l משתנים. רוצים פיענוח מקומי של $p(x)$.

הרעיון: נבחר ישר (תת מרחב אפיני חד מימדי) מקרי דרך הנקודה x $L_z = \{x + t \cdot z | t \in F\}$ (צמצום p לישר הוא ממשתנה אחד t ומאותה דרגה). נגדיר $\phi(t) = p(x + t \cdot z)$. מתוך ערכי p המשובשים על הישר L_z נפענח את הצמצום של p ותשובתנו תהיה $\phi(0) = p(x)$. זאת נעשה ע"י אלגוריתם $Berlekamp Walch$. שראינו. נותר להראות כי בהסברות טובה L_z לא מכיל שיבושים רבים מדי ואת זה נראה בעזרת אי־שיוויון מרקוב. לפני כן נציג משפט.

משפט 5.4 (ללא הוכחה) יהי F שדה סופי, $d, l \in \mathbb{N}$. יש מפענח מקומי לכל קוד RM המטפל ב־$(1 - \frac{d}{|F|})/6$ שיעור טעיות שזמן הריצה שלו הוא $poly(d, l, |F|)$.

ז"א זהו אלג' A בעל זמן ריצה כנ"ל וגישה אקראית ל־$f : F^l \to F$ כך שיש פולינום p ממעלה כוללת $\geq d$ המסכים עם f בהסתברות $\frac{1-\frac{d}{|F|}}{6} - 1$ ומחזיר את p בהסתברות $\leq 2/3$.

נאמר שיש דרך x הוא "מסוכן" אם שיעור הטעויות בו $> (1 - \frac{d}{|F|})/2$. נביט במ"מ המתאים לכל ישר את שיעור הטעויות בו. ע"פ ההנחה

$$\mathbb{E}[x] \leq \frac{1 - \frac{d}{|F|}}{6}$$

ולכן לפי מרקוב,

$$\frac{1}{3} \geq Pr(X > 3 \cdot \mathbb{E}[x])$$

ז"א שבהסתברות לפחות $2/3$הישר המקרי שדגמנו אינו מסוכן ולכן אלג' BW מטפל בבעיה היטב.

6 שיעור 6

6.1 מעגלים בוליאניים

מוגדר מעל n משתנים בוליאניים $x_1, ..., x_n$ וכו'.
נוסיף את הקלטים הקבועים 0 ו 1 ואת $\overline{x_1}, \overline{x_2}, ..., \overline{x_n}$, ונאפשר ליצור צומת חדש בגרף על ידי חיבור שני צמתים קיימים בגרף באמצעות אופרטור $\wedge$ או אופרטור $\vee$.
גודל המעגל יהיה מספר הקשתות.
עומק/גובה המעגל יהיה אורך המסלול הארוך ביותר (הגרף חסר מעגלים ולכן זה מוגדר היטב), ובדרך כלל זה יהיה מקלט לפלט.
אפשר באמצעות מיון טופולוגי, לסדר את הצמתים לפי מרחקם מהקלטים. כל צומת שהוא במרחק $i+1$ מהקלט, אפשר לחשב בזמן קבוע בהנתן ערכי הצמתים במרחק לכל היותר i מהקלט. אנחנו מניחים שהחישוב הוא מקבילי, ולכן עומק המעגל הוא בעצם זמן החישוב המקבילי.
מפרש החוצה/פנימה = $Fan\text{-}In/Out$ הוא חסם על מספר הקשתות הנכנסות יוצאות מקודקוד. במעגלים שהגדרנו מספר הקשתות הנכנסות הוא 2 והיוצאות - בלתי מוגבל.
נטען שלא צריך שלילה כדי לבטא את כל הפונקציות (כלומר לקבל מערכת שלמה).
מדוע?
נניח שאנחנו מאפשרים שערי שלילה, נראה שיש מעגל באותה סיבוכיות שאינו עושה שימוש בשערי שלילה. **הוכחה:** כל צומת של שלילה נחליף בשלילת השרשרת שמובילה אליו (אפשר על ידי החלפת אופרטורים לאורך המסלול עד הקלטים, ואולי בסוף יהיו כל מיני x_k שנחליף ב $\overline{x_k}$ זה מותר. אולי נכפיל את גודל המעגל ב 2, אבל זה לא נורא. ■

הגדרה 6.1 עבור פונקציה בוליאנית $f : \{0,1\}^* \to \{0,1\}$ נגדיר את הקשיות של f להיות:$H(f)(n)$=הגודל המינימלי של מעגל שמחשב את $f(x)$ לכל $x \in \{0,1\}^n$.

טענה 6.2 קיים פולינום $p : \mathbb{N} \to \mathbb{R}^+$ כך שלכל פונקציה בוליאנית $f : \{0,1\}^* \to \{0,1\}$ אם $f \in TIME(t(n))$ אזי לכל n:

$$H(f)(n) \leq p(t(n))$$

הוכחה: בהנתן מכונת טורינג שרצה בזמן $t(n)$ - אזי ניתן לתאר את הסרט הרלבנטי לריצה שלה ע"י רשימה באורך $t(n)$ של סרטים באורך $t(n)$, התא ה i על הסרט ה k "יכיל" מספר חסום של משתנים בוליאניים אשר יאמרו לנו האם הראש של המכונה נמצא בתא הזה, האם האות 0 או 1 כתובה שם וכו'. המשתנה הזה תלוי במספר חסום של משתנים בוליאניים מהסרט ה $k-1$ (כלומר השלב הקודם בריצת המכונה) ולכן נוכל ללא קושי לבנות מכך מעגל בוליאני. כך למשל אם נניח שהמכונה מקבלת כאשר במקום ה 0 כתוב 1 - נבדוק את התא הראשון בסרט האחרון (כלומר את ערך התא הזה בזמן $t(n)$) והוא יכתיב את ערך הפלט של המעגל הבוליאני שנבנה.
כיוון שכל צומת מתקבל ע"י מספר חסום של צמתים מהשכבה הקודמת - נקבל מעגל שגדול פולינומיאלית מ $t(n)$, כנדרש. ■

מה עם הכיוון השני? אם נתון שלכל n מתקיים $H(f)(n) \leq s(n)$, מה נדע מכך על $f \in TIME(?)$?
האמת שלא בהכרח נדע כלום, כי נבחין שבעצם מעגל בוליאני היא סיבוכיות לא אוניפורמית, כלומר לכל n עשוי להיות מעגל שונה, ולכן זה נותן כוח מאד גדול, דומה ל $P_{|poly}$, ולכן f לא בהכרח כריעה.

טענה 6.3 תהי f פונ' בוליאנית כלשהי. $H(f)$ חסומה ע"י פולינום אם"ם $f \in P_{|poly}$.

הוכחה: כקודם - נוכל ליצור מעגל בוליאני שמחקה את האלגוריתם שמובטח מ $P_{|poly}$, ו"לצרוב" בתוכו את הרמז הנכון עבור ה n הספציפי שעבורו אנו בונים את המעגל.
בכיוון השני - אם יש לנו מעגל בוליאני - הרמז יהיה המעגל עצמו... ■

הגדרה 6.4 תהי פונקציה בוליאנית $f : \{0,1\}^n \to \{0,1\}$ נגדיר את

$$d(f) := Minimal\ depth\ of\ a\ boolean\ circuit\ computing\ f$$

.

הגדרה 6.5 יהיו $B_0, B_1 \subseteq \{0,1\}^n$ תת"ק זרות, נסתכל בבעיה הבאה במודל תקשורת לשני שחקנים: עליזה מקבלת $x \in B_0$, בוב מקבל $y \in B_1$. עליזה ובוב צריכים למצוא אינדקס $i \in \{1,...,n\}$ כך ש $x_i \neq y_i$ (יש כזה מכיוון ש x, y שייכים לשתי קבוצות זרות).
נגדיר

$$C(B_0, B_1) = Minimal\ communication\ complexity\ of\ protocol\ to\ agree\ on\ i$$

טענה 6.6 לכל $f : \{0,1\}^n \to \{0,1\}$ מתקיים:

$$d(f) = C\left(f^{-1}(0), f^{-1}(1)\right)$$

למה 6.7

$$C\left(f^{-1}(0), f^{-1}(1)\right) \leq d(f)$$

הוכחה: בה"כ שער הפלט הוא שער $\wedge$, ושני השערים המוליכים אליו מחשבים f_1 ו f_2 כלומר $f = f_1 \wedge f_2$
לעליזה יש $x : f(x) = 0$ ולבוב יש y כך ש $f(y) = 1$ ולכן עליזה קיבלה $f_1(x) = 0$ או $f_2(x) = 0$.
עליזה תעביר ביט לבוב שיספר לו איזה מהביטים הוא 0 (ואם שניהם - אז זה לא משנה איזה תעביר).
נניח ש $f_2(x) = 0$ ועליזה העבירה לבוב מידע זה, וידוע לנו שבוב קיבל 1 בשער ועליסה קיבלה 0, לכן נמשיך באופן רקורסיבי לשאול על השערים ולרדת במורד המעגל, כאשר על כל ירידה ברמה נבזבז ביט אחד של תקשורת, עד אשר נגיע לתא קלט ושנדע שבוב ועליזה לא מסכימים עליו, ולכן זהו ה i המבוקש. ■

למה 6.8 לכל $B_0, B_1 \subseteq \{0,1\}^n$ זרות קיימת $f : \{0,1\}^n \to \{0,1\}$ כך ש:

$$f(B_0) = \{0\},\ f(B_1) = \{1\},\ d(f) \leq C(B_0, B_1)$$

(זה משלים את ההוכחה, כיוון שבפרט אם בהנתן f ניקח את $B_0 = f^{-1}(0)$ ואת $B_1 = f^{-1}(1)$ נקבל את המבוקש.
הוכחה: (באינדוקציה על $C(B_0, B_1)$)
למשל עבור $0 = C(B_0, B_1)$ (נניח אם הקבוצות הן כל הקלטים שבהן הקוא' ה 7 היא אפס וכל הקלטים שבהן הקוא' ה 7 היא 1, ואז הפרוטוקול משיב 7 באופן קבוע)
במקרה כזה, אם הפרוטוקול פולט תמיד i, נבנה מעגל שבו הפלט הוא בדיוק $\overline{x_i}$.
צעד האינדוקציה:
נניח שהוכחנו עבור פרוטוקול באורך k, ונראה עבור פרוט' באורך $k+1$.

נניח בה"כ שעליזה מדברת קודם.
נסמן ב $B_0^0, B_0^1 \subseteq B_0$ את הקלטים עבורם הביט הראשון שעליזה שולחת הוא $0,1$ בהתאמה.
הבחנה ש $C\left(B_0^i, B_1\right) \leq C\left(B_0, B_1\right) - 1$ לכל i.
מדוע?
נדע שההודעה הראשונה שנשלחה היא i, ולכן לא צריך לשלוח את הביט הזה, ונסמלץ את שאר הפרוטוקול כרגיל.
מכאן שיש מעגל בעומק $C\left(B_0^0, B_1\right)$ עבור פונקציה f_0כלשהי, המקיימת $f_0\left(B_0^0\right) = \{0\}$ ו $f_0\left(B_1\right) = \{1\}$
ויש מעגל בעומק $C\left(B_0^1, B_1\right)$ עבור פונקציה f_1 המקיימת $f_1\left(B_0^1\right) = \{0\}$ ו $f_1\left(B_1\right) = \{1\}$
כעת מתקיים $f = f_0 \wedge f_1$ ואם נוסיף צומת $\wedge$ שנחבר אליו את פלטי שני המעגלים - נקבל שהעמקנו את המעגלים הקודמים ב 1 לכל היותר, ומההבחנה נובע המבוקש. ■

6.2 קשיות והגברתה

הגדרה 6.9 בהנתן $f : \{0,1\}^* \to \{0,1\}$ ו α. נגדיר את $H_\alpha(f)(n)$ להיות הגודל המינימלי של מעגל שמחשב את f על לפחות החלק ה α מהקלטים $x \in \{0,1\}^n$.

כך למשל:

$$H_1(f)(n) = H(f)(n)$$

$$H_{\frac{1}{2}}(f)(n) \leq 1$$

משום שבהכרח יש ניחוש לתוצאת הפונקציה שנכון בלפחות מחצית מהפעמים.
יכול להיות שבשביל להחליט איזה מעגל נבחר (כזה שמנחש תמיד 0 או תמיד 1) נצטרך לעבוד המון, אבל המעגל עצמו יהיה קטן.
דרך להוכיח למשל את $NP \neq P$ היא להראות ש $NP \nsubseteq P_{|poly}$, כלומר למצוא $f \in NP$ כך ש $H(f)$ גדול אסימפטוטית מכל פולינום.
למצוא $f \in NP$ כך ש $H(f)(n) > 20n$ זו בעיה פתוחה...
כך שמצבנו קשה מאד, והמרחק למצוא פונקציה מפורשת שהקשיות שלה היא גדולה מכל פולינום - עוד רב.
לא יודעים האם $EXP \subseteq P_{|poly}$

טענה 6.10 תהי $s : \mathbb{N} \to \mathbb{N}$ ונניח שיש $f \in EXP$ כך ש $s(n) \leq H_1(f)(n)$
אזי יש $g \in EXP$ וקבוע $c > 0$ כך ש:

$$\frac{s\left(\frac{n}{c}\right)}{n^c} \leq H_{0.99}(g)(n)$$

הוכחה: אם נתבונן בטבלת האמת של f עבור קלטים בגודל n בתור מחרוזת באורך 2^n.
נתבונן ב g כקוד ריד-מולר של f (פחות או יותר) מורכב עם הדמר (כדי שהפלט יהיה בוליאני ולא איברים בשדה, שזה מה שריד מולר פולט), ונתבונן ב g כמחרוזת.
נקבל משהו באורך N^5 עם יכולת לתיקון של יותר מאחוז אחד מהשגיאות, ולכן בהנתן מעגל שמחשב את g נכונה ב 99% מהביטים - נפעיל את אלגוריתם תיקון השגיאות של ריד סולומון כדי למצוא את המחרוזת שמייצגת את טבלת האמת של f, בהסתברות של 90% נתקן את השגיאות נכונה.
זה עדיין קצת בעייתי כי הצגנו אלגוריתם ולא מעגל, וכן אנחנו רוצים לפתור ללא שגיאה, אבל התמודדנו עם בעיות דומות בעבר (היום ובתרגיל הבית), נמשיך בשיעור הבא להציג בדיוק כיצד נעשה זאת. ■

7 שיעור 7

נמשיך את הוכחת הטענה האחרונה מהשיעור הקודם.
מה שנרצה להראות זה שאם מניחים קיום של f קשה אקספוננציאלית, כלומר שדרוש זמן אקספוננציאלי כדי לחשב אותה, נוכל להראות קיום g שדרוש זמן אקספוננציאלי כדי לחשב אותה גם עבור 0.99 מהקלטים.
ננסח זאת שוב לצרכי בהירות:

משפט 7.1 אם יש $f \in ExpTime$ כך ש $H_1(f)(n) \geq s(n)$
אז קיים קבוע k כלשהו ויש $g \in ExpTime$ כך ש $H_{0.99}(g)(m) \geq \frac{s\left(\frac{m}{k}\right)}{poly(m)}$

נבנה את g כך שיהיה קוד תיקון שגיאות עבור f.
עד כה ראינו קודי $RM(\mathbb{F}, l, d)$ כך ש $\mathbb{F}$ השדה, l המימד ו d הדרגה.
כלומר אוספי מילים מעל הא"ב $\mathbb{F}$ שמקודדות את ערכיו של פולינום מעל $\mathbb{F}^l$ ודרגתו קטנה (שווה) מ d.
במילים אחרות - אם יש לנו פולינום מדרגה d - אז מילת קוד חוקית מתאימה בדיוק למטריצה l מימדית המכילה את ערכי הפולינום בכל הנקודות ב $\mathbb{F}^l$. **הוכחה:** תהי $f : [N] \to \underbrace{\{0,1\}}_{\mathbb{F}}$ ונרצה לקודד אותה כמילה $h : [M] \to \mathbb{F}$
בזמן פולינומי ב N, כך שהערכים של f יופיעו במקומות מוגדרים כערכים של h. למה הכוונה? שאם נרצה לחשב נקודה מסויימת ב f - תהיה נקודה מסויימת ב h שמספיק לחשב אותה כדי לקבל תוצאה זו. (נזכיר כי $N = 2^n$ כלומר f היא טבלת אמת עבור מחרוזות באורך n)
נשתמש בקוד $RM(\mathbb{F}, l, d)$, מילת קוד תהיה פולינום ב l משתנים מדרגה d מעל השדה $\mathbb{F}$,
כאשר $|\mathbb{F}| = \log^5 N$ (הגודל 5 לא חשוב, אבל נרצה שהוא יהיה פולילוגריתמי ב N)
ו $l = \frac{\log N}{\log\log N}$ (זה חשוב כדי להשיג זמן פולינומי)
$d = \log^2 N$
אורך מילה יהיה למעשה מספר תאי המטריצה ה l מימדית, כלומר $M = |\mathbb{F}|^l = \left(\log^5 n\right)^{\frac{\log N}{\log\log N}} = N^5$ שהוא פולינומיאלי ב N.
נזכיר כי רצינו שערכי f יהיו ממש תת קבוצה של ערכי g, ולכן ניקח את המטריצה ה l מימדית שנבנה, נגדיר $H \subseteq \mathbb{F}$ כלשהי כך ש $|H| = \frac{d}{l}$ ולכן תהיה קוביה H^l שערכיה יהיו ערכי f.
נבחר $H \subseteq \mathbb{F}$ (לא נסביר איך בוחרים, כי זה טכני ולא רלבנטי) בגודל $|H| = \frac{d}{l}$ ונשכן את $[N]$ בתוך הקבוצה H^l, איך נשכן? לא נראה אלגוריתם ספציפי, ונניח שיש לנו $\phi : [N] \to H^l$ חח"ע שפועלת בזמן פולינומי וממפה את $[N]$ לתת־מטריצה ה l מימדית H^l.
זה אפשרי מבחינת שיקולי גודל, שכן

$$\left|H^l\right| = |H|^l = \left(\frac{d}{l}\right)^l = (\log N \cdot \log\log N)^{\frac{\log N}{\log\log N}} > N$$

ולכן בקוביה H^l יש "מספיק מקום" כדי להכיל את טבלת האמת של f.
עתה נמצא פולינום $h : \mathbb{F}^l \to \mathbb{F}$ שדרגתו לכל היותר d, אשר ב H^l מסכים עם השיכון שלנו. דהיינו:

$$\forall x \in [N] : H(\phi(x)) = f(x)$$

בעצם h שנמצא דרגתו בכל משתנה תהיה לכל היותר $\frac{d}{l}$ (וזה חזק יותר ממה שאנחנו צריכים וגורר אותו)
מספיק למצוא לכל $x_0 \in H^l$ פולינום h_{x_0} כנ"ל (כלומר שדרגתו בכל משתנה וכו') המקיים:

$$h_{x_0}(x) = \begin{cases} 1 & x = x_0 \\ 0 & x \neq x_0 \end{cases}$$

ואז h שמחפשים יוגדר להיות:

$$h = \sum_{x_0 \in H^l \,:\, f(\phi^{-1}(x_0))} h_{x_0}$$

כלומר לקחת את כל הנקודות שאנחנו מזהים עם ערכי f, ולסכום את הפונקציות h_{x_0} שמתקבלות בנקודות הללו.

אז איך נגדיר את h_{x_0}? (בנוסחה הבאה $x^{(i)}$ משמעות הקוא' ה i באיבר x, נבחין כי האיברים במטריצה ה l מימדית שלנו הם תאים המתוייגים ע"י l קוא' כל אחד)

$$h_{x_0}(x)=\prod_{i=1}^{l}\prod_{\alpha\in H\setminus\left\{x_0^{(i)}\right\}}\left(x^{(i)}-\alpha\right)$$

כרגע זה באמת מתאפס לכל $x\neq x_0$, ננרמל כדי שזה יתן 1 על x_0 ונקבל:

$$h_{x_0}(x)=\frac{\prod_{i=1}^{l}\prod_{\alpha\in H\setminus\left\{x_0^{(i)}\right\}}\left(x^{(i)}-\alpha\right)}{\prod\prod\left(x_0^{(i)}-\alpha\right)}$$

מה דרגת הפולינום?

נבחין כי עבור $i=7$ למשל - המשתנה $x^{(7)}$ מופיע רק כאשר במכפלה החיצונית יש $i=7$, במכפלה הפנימית יש $|H|-1$ איברים וזה פחות מ $\frac{d}{l}$.

כיוון שביצירת h אנו סוכמים פולינומים שדרגתם בכל משתנה קטנה מ $\frac{d}{l}$ - נקבל שהפולינום הכולל הוא כזה גם כן, כפי שרצינו.

זמן החישוב גם הוא פולינומי, כיוון שבכל h_{x_0} אנחנו מחשבים "לא הרבה איברים" במונה ובמכנה, וכדי לחשב את h הכללית צריך לסכום N פונקציות h_{x_0} שונות, ולכן בסך הכל זה יצא פולינומי.

כעת - מחישוב מקורב של h לחישוב מדוייק של f

מה מרחק הקוד? לפי שוורץ זיפל שני פולינומים מדרגה d ישתוו על לא יותר מהחלק ה $\frac{d}{\mathbb{F}}$ של הקלטים, ולכן המרחק של הקוד יהיה $1-\frac{d}{\mathbb{F}}=1-\frac{1}{Poly\log(N)}$

נגדיר את m,n כך ש $2^m=M$ ו $2^n=N$.

כעת אפשר לחשוב על f ו g כפונקציות ממחרוזות באורך n ו m כלומר:

$$f:\{0,1\}^n\to\{0,1\}\quad g:\{0,1\}^m\to\{0,1\}$$

טענה 7.2 אם יש מעגל C בגודל $s(m)$ המחשב את h על $1-\frac{1}{6}\left(1-\frac{d}{\mathbb{F}}\right)$ מהקלטים, אזי יש C' רנדומי (שיש לו אפשרות לקבל כחלק מהקלט שלו מחרוזת רנדומית שאורכה פולינומי ב n) בגודל $s(m)\cdot poly(m)$ כך שלכל x:

$$Pr\left[C'(x)=f(x)\right]\geq\frac{2}{3}$$

אז מה זה בעצם $s(m)\cdot poly(m)$? נגדיר את זה להיות $s'(n)$:

$$s'(n)=s(m)\cdot poly(m)=s(5n)\,poly(n)$$

או אז נקבל $s(m)=\frac{s'\left(\frac{m}{5}\right)}{poly(m)}$

הוכחה: נתבונן ב C, בגלל ש h מקבלת ערכים בשדה - אז יש לנו כמה שערי פלט שנתרגם לקלט ב $\mathbb{F}$, אבל זה לא באמת משנה. על פי הנחתנו על חמש שישיות מהקלטים המעגל משיב נכונה.

אנו רוצים לחשב את $f(x)$ בהסתברות טובה.

לא מספיק לנו לחשב את $C(\phi(x))$ שכן זה אמנם C יעבוד באופן כללי בהסתברות טובה על כל $\mathbb{F}^l$, אבל אולי C ממש טועה על כל איברי H^l ...

מה עושים?

משתמשים בריד מולר, ועושים בו פיענוח מקומי. לוקחים את התא המבוקש מתוך H^l, שאת ערכו נבקש לחשב, נעביר ישר מקרי במטריצה $\mathbb{F}^l$ שעובר דרך התא המבוקש, וראינו שבהסתברות טובה יתקבל קידוד ריד מולר שבו בתא המבוקש יש את הערך המתאים.

מהידוע לנו על פיענוח לוקאלי נקבל:

$$Pr\left[C'\left(x\right)=f\left(x\right)\right]\geq\frac{2}{3}$$

ולכן נותר לבדוק כמה משאבים אנחנו מבזבזים:
כמה נקודות יש על ישר בעולם $\mathbb{F}^l$? נוסחת הישר תהיה $a+t\cdot b$ כאשר t רץ על כל איברי $\mathbb{F}$, ואת a ו b נבחר כך שהישר יעבור בתא שלנו, מכאן שבישר יש $\mathbb{F}$ נקודות.
מכאן שנעריך את C ב $\mathbb{F}$ ערכים, כלומר במספר פולילוגריתמי ב N, וזה פולינומי ב n וגם ב m, בנוסף חישוב ריד מולר יעלה לנו משהו פולינומי באורך המילה ולכן סה"כ:

$$\left|C'\right|\leq\left|C\right|\cdot poly\left(m\right)$$

■

טענה 7.3 קיים קבוע אוניברסלי k כך שאם $f:\{0,1\}^n\to\{0,1\}$ ניתנת לחישוב בזמן אקספוננציאלי, אז יש $h:\{0,1\}^m\to\{0,1\}$ שניתנת לחישוב גם כן בזמן אקספ', וכך שאם מעגל C, בגודל $s\left(m\right)$, מחשב את h על 0.9 מהקלטים בגודל m, אז יש מעגל C' המחשב את f על קלטים בגודל $\frac{m}{k}$ ומתקיים $\left|C'\right|\leq m\cdot poly\left(m\right)$

מסקנה 7.4 אם $f'\in ExpTime$ מקיימת $H_1\left(f\right)\left(n\right)\geq s\left(n\right)$ אזי יש $h\in ExpTime$ כך ש־:

$$H_{0.9}\left(h\right)\left(m\right)\geq\frac{s\left(\frac{m}{k}\right)}{poly\left(m\right)}$$

הבעיה היחידה שנותרה היא ש h אינה פונקציה בוליאנית. שכן:

$$h:\{0,1\}^m\to\{0,1\}^{\log\log m}$$

איך נבנה את g הבוליאנית שלנו? נמיר את הקוד לבינארי ונחליף כל "מילה בינארית" שמייצגת "אות" בשדה $\mathbb{F}$ בקוד הדמר שלה, ונתרגם אותה בהתאם, ואז $g\left(x,i\right)=Had\left(h\left(x\right)\right)^{(i)}$ ו:

$$g:\{0,1\}^{m+\log m}\to\{0,1\}$$

(מדוע אין אנו מסתפקים בתרגום בינארי? כיוון שאז יהיה אפשר להחליף מספר לוגריתמי של ביטים אבל לפגום בכל המילה, וקוד הדמר שומר שכל עוד השתבשו פחות ממחצית מהביטים ב"מילה הבינארית" נדע לתקן זאת, שכן קוד הדמר מנפח את הקידוד שלנו, שהוא $\log\log N$ באופן אקספוננציאלי) ■

8 שיעור 8

8.1 תיקון מקומי - תזכורת

מהו אלגוריתם לפיענוח מקומי של קוד $E:\mathbb{F}^n\to\mathbb{F}^m$ הפועל במרחק מנורמל ρ?
זהו אלגוריתם D הפועל בהנחות הבאות: יש $y\in\mathbb{F}^n$ כך שיש $x\in\mathbb{F}^n$ יחיד המקיים $\frac{d_H}{m}\left(y,E\left(x\right)\right)<\rho$ (זהו מרחק המינג מנורמל), האלגוריתם D מקבל גישה אקראית ל y ואינדקס j ומחשב את x_j (הקוא' ה j של x) בהסתברות לא טריוויאלית (כלומר חסום מ $\frac{1}{\mathbb{F}}$, למשל אם $\mathbb{F}=\{0,1\}$ אז $\frac{2}{3}$).
הצורה האופיינית שבה יתממשו כאן הרעיונות הללו: יש פולינום P לא ידוע ("מילת קוד", וזהו $E\left(x\right)$) מוכרת לנו גרסה מורעשת (=מקולקלת) של P, היא f, ע"י גישה מקרית ל f. נרצה לדעת את $P\left(\underbrace{x^*}_{j}\right)$, כלומר לשחזר את ערך הפולינום הבלתי מורעש בנקודה ספציפית.

מדוע זה מועיל לפיענוח? כיוון שראינו שאפשר לדרוש שהמילה אחרי הקידוד, תכיל את כל הקואורדינטות של המילה שלפני הקידוד, אולי לא באותו הסדר ולא ברצף, אבל בתבנית ידועה, ולכן בהנתן האלגוריתם D, נוכל ממש לפענח.

כזכור, ראינו אלגוריתם פולינומי כנ"ל לפיענוח מקומי של קוד RM:

1. פיתחנו אלגוריתם יעיל לפיענוח של RS, מחפשים פולינומים שמתאפסים במקומות מתאימים ופותרים משוואות ליניאריות, $Berlkamp\,Walch$.
2. האלגוריתם עצמו: בהנתן לנו גישה אקראית לפונקציה $f:\mathbb{F}^w\to\mathbb{F}$ שקרובה דיה לפולינום ממעלה $d\geq$ ב n משתנים, ובהנתן לנו $x\in\mathbb{F}^l$ רצינו לדעת בהסתברות לא אפסית מהו $P(x)$. דרך הנקודה $x\in\mathbb{F}^l$ מעבירים ישר מקרי L, הצמצום של P ל L, כלומר נבחר באקראי $b\in\mathbb{F}^l$ ואז הישר יהיה $L=\{x+bt:t\in\mathbb{F}\}$ במשתנה היחיד $t\in\mathbb{F}$. את הצמצום הנ"ל אנחנו רואים כקוד RS ואותו אנו מפענחים בעזרת אלגוריתם $BW\Rightarrow Berlkamp\,Walch$ ותשובתנו היא $Q(0)$.

היום נעבוד בשלבים הבאים:

1. פיענוח רשימות מקומי
2. אלגוריתם Sudan לפיענוח רשימות של RS
3. דרך x נעביר עקום ממעלה (שניה) שלישית y נצמצם את RM ל y, גם כאן יתקבל קוד RS בפרמטר $t\in\mathbb{F}$ על הצמצום נפעיל את אלגוריתם Sudan ונגמור כמקודם.

8.2 פיענוח רשימות של RS (M. Sudan 96')

אם $E:\mathbb{F}^n\to\mathbb{F}^m$ קוד לתיקון שגיאות, אז בפיענוח רשימות שלו נרצה, בהנתן לנו $y\in\mathbb{F}^m$ לדעת את כל מילות הקוד בסביבה מרדיוס נתון סביב y.

משפט 8.1 יש אלגוריתם יעיל שבהנתן לו $(a_1,b_1),...,(a_m,b_m)\in\mathbb{F}^2$ מחזיר את כל הפולינומים G במשתנה יחיד ממעלה $d\geq$ כך ש $G(a_i)=b_i$ לפחות עבור $2\sqrt{dm}$ ערכים של i.

הוכחה: נמצא פולינום $Q(x,y)$ שמעלתו ב x היא $\sqrt{dm}\geq$ ומעלתו ב y היא $\sqrt{\frac{m}{d}}\geq$ כך שלכל i מתקיים $Q(b_i,a_i)=0$.

נתרגם את הדרישה על Q למערכת של משוואות ליניאריות הומוגניות במקדמי Q.

מספר המשוואות הוא m ומספר המשתנים הוא $\left(1+\sqrt{dm}\right)\left(1+\sqrt{\frac{m}{d}}\right)>m$, וכיוון שכך יש לפחות פתרון אחד לא טריוויאלי, כלומר יש Q כנ"ל.

(מדוע זה מספר המשתנים? שכן עבור

$$Q=\sum_{0\leq i\leq\alpha\wedge 0\leq j\leq\beta}a_{ij}x^iy^j$$

יש $(1+\alpha)(1+\beta)$ מקדמים)

יש אלגוריתמים יעילים לפירוק לגורמים, ובעזרתם נפרק את Q לגורמיו. (לא נפרט על כך)

(*) לכל גורם G של Q מהצורה $G(x,y)=y-P(x)$כשמעלת P היא $d\geq$ ו $P(a_i)=b_i$ לפחות ל t ערכי i נקבל את P לרשימה.

נזכיר שאנחנו מקבלים כקלט רשימה של m נקודות וערכיהן בפולינום ורוצים כפלט לתת רשימה של כל הפולינומים P במשתנה יחיד, המועמדים לתפקיד הפולינום המקורי, שמעלתם $d\geq$, וכך שלפחות עבור t מה i-ים (כאשר $t=2\sqrt{md}$) מתקיים $P(a_i)=b_i$.

בכיוון אחד - ברור שכל פולינום שיתקבל על ידי האלגוריתם, לפי דרישת (*) מקיים את הנחוץ לנו.

בכיוון השני - יש להוכיח שאם Φ עומד בדרישות המשפט, ז"א פולינום במשתנה יחיד ממעלה $d\geq$ המקיים $\Phi(a_i)=b_i$ לפחות t פעמים, אז $y-\Phi(x)$ הוא גורם בפירוק של Q.

נביט בפולינום $Q(x,\Phi(x))$, וזה פולינום במשתנה יחיד x, ומעלתו $\geq$

$$t>2\sqrt{md}=\underbrace{d\cdot\sqrt{\frac{m}{d}}}_{deg\,Q\,times\,deg\,\Phi}+\underbrace{\sqrt{md}}_{deg\,Q\,in\,x}$$

והפולינום הזה מתאפס t פעמים ולכן הוא פולינום ה 0

$$Q(x,\Phi(x))\equiv 0\Rightarrow Q|\,(y-\Phi(x))$$

נסביר את המעבר האחרון:
כזכור מאלגברה לינארית, הפולינום $A(z)$ מתחלק ב $z-\alpha$ אם"ם $A(\alpha)=0$.
זה אנלוגי למה שאנחנו רוצים להוכיח, כלומר לקחת את $Q(x)$ כשדה הפונקציות הרציונליות ב x, ולהתבונן בחוג הפולינומים ב y מעליו, ואז $\Phi(x)$ הוא איבר בשדה ולכן הפולינום $Q(x,y)$ (שהוא בעצם פולינום במשתנה y) מתחלק ב $y-\Phi(x)$ אם"ם הפולינום מתאפס עבור הצבת $\Phi(x)$, ואכן ראינו שכך הוא, כלומר במקרה שלנו - התוצאה היא שמקבלים זהותית את פולינום האפס.
וכיוון שהראינו את שני הכיוונים - קיבלנו שהאלגוריתם אכן משיג את מטרתו המוצהרת. ■

כעת נרצה לבחון את אורך הרשימה המתקבל בפיענוח הרשימות ולהראות שהוא לא גדול מדי. בפרט מההוכחה הקודמת, כיוון שכל פולינום שמצאנו היה מתוך פירוק של Q, מספר הפולינום לא גדול מדרגת Q.

טענה 8.2 תהי $f:\mathbb{F}\to\mathbb{F}$ כלשהי, יהי $d\geq 0$ שלם ו $\varepsilon>2\sqrt{\frac{d}{|\mathbb{F}|}}$ אז יש לא יותר מ $\frac{2}{\varepsilon}$ פולינומים ממעלה $\geq d$ אשר מתלכדים עם f בלפחות $\varepsilon\cdot|\mathbb{F}|$ מנקודות השדה.

הוכחה: יהיו $P_1,...,P_k$ כל הפולינומים כנ"ל, נוכיח ש $k\leq\frac{2}{\varepsilon}$
נחשוב על $\mathbb{F}$ כמרחב הסתברות בהתפלגות אחידה, ונביט במאורע $A_i\subseteq\mathbb{F}$ המוגדר :

$$A_i=\{x\in\mathbb{F}:P_i(x)=f(x)\}$$

הערה: כאשר קוטעים את נוסחת ההכלה וההדחה במקום זוגי מקבלים חסם תחתון (ובמקום אי־זוגי מקבלים חסם עליון) על ההסתברות של האיחוד, ולכן:

$$1\geq Pr\left(\bigcup A_i\right)\geq\sum Pr(A_i)-\sum_{i<j}Pr(A_i\cap A_j)$$

נבחין כי בסכום הראשון כל מחובר גדלו לכל היותר $\frac{d}{|\mathbb{F}|}$, וכל מחובר בסכום האחרון הוא בעצם חיתוך של שני מאורעות כנ"ל, כלומר מספר הנקודות שעליהם מסכימים שני פולינומים מדרגה לכל היותר d וזה כמובן d לכל היותר, ומכאן:

$$1\geq k\cdot\varepsilon-\binom{k}{2}\frac{d}{|\mathbb{F}|}$$

אם ניקח $k=\frac{2}{\varepsilon}$ נקבל

$$1\leq\frac{d}{|\mathbb{F}|}\binom{\frac{2}{\varepsilon}}{2}$$

$$\frac{2}{\varepsilon^2}\cdot\frac{d}{|\mathbb{F}|}>1$$

$$\frac{2d}{|\mathbb{F}|}>\varepsilon^2\Rightarrow\sqrt{\frac{2d}{|\mathbb{F}|}}>\varepsilon$$

בסתירה להנחתנו על בחירת אפסילון, ומכאן ש $k<\frac{2}{\varepsilon}$ כנדרש, ובעצם אין "המון" פולינומים המסכימים עם f תחת ההנחות הנ"ל. ■

8.3 פיענוח רשימות מקומי

יש f קרובה ל P, יש לנו גישה אקראית ל f, בקלט x רוצים את $P(x)$ בהסתברות הצלחה לא טריוויאלית.
היינו רוצים למצוא בסיכוי הצלחה לא טריוויאלי את $P^*(x)$ כאשר P^* הוא פולינום אחד מסויים מבין אלה הקרובים ל f. הבעיה היא לתת אפיון (מציין) ל P^* שבו אנו מעוניינים אף כי הרשימה איננה לפנינו.
מדוע? כיוון שאחרת היינו מקבלים בכל פעם (אולי) פולינום אחר, שהוא בעצם מילת קוד אחרת, והפיענוח שלנו היה משתנה לפי האינדקס שאליו אנחנו ניגשים, וזו קטסטרופה אם אנחנו מצפים ממנו להיות קונסיסטנטי.
איך ניתן "להצביע" על P^*? למשל $P(x_0)=y_0$

הגדרה 8.3 יהי $E:\{0,1\}^n\to\{0,1\}^m$ קוד מתקן שגיאות ו $\rho=1-\varepsilon$ ($\varepsilon>0$) אומרים ש D הוא פיענוח רשימות מקומי ל E המטפל ב ρ שגיאות אם לכל $y\in\{0,1\}^m$ ו $x\in\{0,1\}^n$ המקיימים $\Delta(E(x),y)\le\rho$ יש אינדקס $i_0\in\left[poly\left(\frac{n}{\varepsilon}\right)\right]$ כך שלכל $j\in[n]$ בהנתן קלטים i_0,j וגישה מקרית ל y, האלגוריתם D רץ בזמן $poly\left(\frac{\log m}{\varepsilon}\right)$ ופולט את x_j בהסתברות $\ge\frac{2}{3}$.

(כאשר x הוא האיבר ה i_0 ברשימת איברי $\{0,1\}^n$ המקיימים את התנאי $\Delta(E(x),y)\le\rho$.)

משפט 8.4 לקוד RM יש מפענח רשימות מקומי המטפל ב $1-10\sqrt{\frac{d}{|\mathbb{F}|}}$ שגיאות. ז"א לכל $\mathbb{F},d,l$ יש אלגוריתם D בעל זמן ריצה $poly(\mathbb{F},d,l)$, לאלגוריתם יש גישה מקרית לפונקציה $f:\mathbb{F}^l\to\mathbb{F}$ והוא מקבל שני קלטים $x\in\mathbb{F}^l$ ו $i\in\mathbb{F}^{l+1}$ ופולט איבר ב $\mathbb{F}$.
אם הפונקציה fמתלכדת עם פולינום כלשהו $P:\mathbb{F}^l\to\mathbb{F}$ ממעלה $\ge d$ על לפחות $10\sqrt{\frac{d}{|\mathbb{F}|}}$ מהנקודות, אז יש $i_0\in\mathbb{F}^{l+1}$ כך שמתקיים

$$Pr\left(D^f(i_0,x)=P(x)\right)\ge\frac{2}{3}$$

על פני הבחירות האקראיות של D

הטענה במשפט חלה על כל x!
אנחנו נוכיח זאת רק עבור 0.9 מה xים, ואז נפעיל את אלגוריתם הפיענוח של RM הרגיל, על מנת להשיג זאת לכל x.
הערה: אנחנו חושבים על $i_0\in\mathbb{F}^{l+1}$ בתור זוג $(x_0,y_0)\in\mathbb{F}^l\times\mathbb{F}$

9 שיעור 9

טענה 9.1 נניח ש $f\in Exp$ מקיימת $H(f)(n)\ge s(n)$ עבור $poly(n)<s(n)<2^n$
אזי יש פונקציה $g\in Exp$ וקבוע $C>0$ כך ש $H_{avg}(g)(m)\ge s\left(\frac{m}{c}\right)^{\frac{1}{c}}$ החל מ $m>m_0$ כלשהו.

נזכיר ש $H_{avg}(g)(m)>s$ משמעו $H_{\frac{1}{2}+\frac{1}{s}}(g)(m)>s$

הגדרה 9.2 אם $\psi:\Sigma_1^N\to\Sigma_2^M$ קידוד. אז מפענח רשימות מקומי מתאים המבצע q שאילתות ועובד עבור רשימות בגודל L ומרחק δ (אפשר לומר באופן דומה הסכמה $\alpha=1-\delta$) הוא מעגל/אלגוריתם הרץ בזמן q ומקיים שלכל $y\in\Sigma_2^M$ (כלומר מילה במרחב שבו חיות מילות הקוד) ולכל $x\in\Sigma_1^N$ כך ש $d(\psi(x),y)<\delta$ (וכאן באופן דומה אפשר לומר $agr(\psi(x),y)>\alpha=1-\delta$) קיים $t\in\{1,...,L\}$ (אינדקס לאיבר ברשימה) עבורו מתקיים:

$$\forall j\in\{1,...,N\}\ C(y,j,t)=x_j$$

הוכחה: (סכמה של הוכחה לטענה)
בהנתן $f:\{0,1\}^n\to\{0,1\}$ כמו בשאלה, נתבונן ב"מילה" שהיא טבלת האמת שלה, ואורך המילה לכן יהיה $N=2^n$.
נייצר $g:\{0,1\}^{kn}\to\{0,1\}$ שטבלת האמת שלה תהיה באורך $M=2^{k\cdot n}$.

כלומר הקידוד יהיה:

$$\psi : \{0,1\}^N \to \{0,1\}^M$$

ולכן $g = \psi(f)$.

קל לראות ש g אכן ניתנת לחישוב בזמן אקספוננציאלי (מנימוקים שכבר ראינו בעבר).

נבחר $q = s(n)^\varepsilon$ ואז בהנתן מילה $h \in \{0,1\}^M$ שמסכימה עם g על $\frac{1}{2} + \frac{1}{s(n)^3}$ אז נציג מעגל C המקיים $C(h,x,t) = f(x)$ (x כאן הוא אינדקס לתא במילה f, כלומר על תקן j).

C רץ בזמן פולילוגריתמי ב N, כלומר בזמן פולינומי ב n, ומחזיר את $f(x)$.

מה זה אומר? שאם h קלה לחישוב, אזי לא ניתן למעגל שלנו את h בייצוגה המלא (כי ממילא זמן הריצה שלו לא יספיק כדי לקרוא את כולה), אלא רק גישה למעגל פולינומי שמחשב את h, ואז C שלנו בזמן פולינומי יכול לחשב את $f(x)$. ■

כלומר אם נציג קוד תיקון שגיאות בפרמטרים שהוצגו (הדרישה הרצינית פה היא ש q פולילוגריתמי באורך המילה, כלומר אלגוריתם הפיענוח קצר מאד, ועדיין מצליח לחשב ביט אחד מתוך הקלט).

9.0.1 הקידוד

<u>הקידוד שנשתמש בו:</u>

נהפוך את טבלת האמת של f (שתאיה מתוך $\{0,1\}$) מילה g' שאותיותיה ב $\mathbb{F}$ כלשהו, ובאמצעות קוד WH נהפוך מילה זו ל (f) $g = WH \circ RM\left(\underbrace{s(n)^\varepsilon}_{|\mathbb{F}|=field\,size}, \underbrace{s(n)^{\frac{\varepsilon}{2}}}_{d=degree}, \underbrace{\frac{2n}{\log d}}_{l=dimension}\right)$ ונקבל מילה שאותיותיה ב $\{0,1\}$.

9.0.2 פיענוח

איך נפענח הרכבה של הקודים? נפעיל את אלגוריתם הפיענוח של RM וכאשר נרצה את הערך בתא מסויים של g' (ונתונה לנו רק g) נחשב את הפיענוח של WHעבור המקום הזה. ביתר פירוט:

נניח ש $\psi : \{0,1\} \to \Sigma^{M'}$ מגיע עם פיענוח רשימות לוקלי עם q_1 שאילתות, הסכמה α ואורך רשימות L_1.

ו $\Phi : \Sigma^{M'} \to \{0,1\}^M$ עם פיענוח רשימות לוקלי עם q_2 שאילתות, הסכמה $\frac{1}{2} + \varepsilon$ ואורך רשימות L_2.

טענה 9.3 ל $\Phi \circ \psi$ יש פיענוח רשימות לוקלי עם $q_1 \cdot q_2$ שאילתות, הסכמה $\frac{1}{2} + \varepsilon + \alpha \cdot L_2$ ואורך רשימות $L_1 \cdot L_2$.

הוכחה: אם C מפענח רשימות המתאים ל ψ ו D מפענח רשימות ל Φ, נבנה מפענח E להרכבתן :

$$E(h,x,t_1,t_2)_{t_1 \in \{1,\ldots,L_1\}, t_2 \in \{1,\ldots,L_2\}}$$

ו Eיפעיל את $C(D(h,\cdot,t_2),x,t_1)$, זה קצת abuse of notation אבל הכוונה היא שלא ניתן למעגל את h במלואה, אלא גישה ל D עם הפרמטרים המתאימים, וזה שקול למתן $D(h)$ במפורש.

טענה: נניח של h יש הסכמה $\frac{1}{2} + \varepsilon + \alpha L_2$ עם הקידוד $\Phi \circ \psi(f)$, אזי עבור לפחות αL_2 מהאיברים y מתקיים (כאן y הוא התא ה y בקידוד של f ע"י ψ):

$$agr\left(\Phi(\psi(f)(y)), h_{|y}\right) \geq \frac{1}{2} + \varepsilon$$

הוכחה לטענה: מא"ש מרקוב (דילגנו על להראות את זה במדויק)

לכן לפחות על αL_2 מהאיברים יש הסכמה $\varepsilon + \frac{1}{2}$, ולכן יש t_2 שכאשר נשתמש בו - נפענח את $\psi(f) = g'$ על לפחות α מהקלטים.

ומכאן שעבור t_2זה יש t_1 שעבורו המפענח יחשב את $f(x)$ נכונה לכל x. ■

9.1 מעגלים ופסאודו רנדומיות

תהי D התפלגות על קלטים מ $\{0,1\}^n$, נאמר ש D היא ε־פסאודו־רנדומית למעגלים בגודל s אם לכל מעגל C בגודל $s \geq$ מתקיים

$$|P_{x\sim D}\left(C\left(x\right)=1\right)-P_{x\sim U_n}\left(C\left(x\right)=1\right)|<\varepsilon$$

בהנתן פונקציה $g:\{0,1\}^n\to\{0,1\}^m$ (נחשוב ש m גדול מ n) נאמר ש־g היא ε־יוצר פסאודו רנדומי אם למעגלים בגודל s, $\{g\left(x\right)\}_{x\sim U_n}$היא התפלגות ε־פסאודו רנדומית עבור מעגלים בגודל s.

אבחנה: עבור s פולינום ב n יש $g:\{0,1\}^n\to\{0,1\}^m$ שהיא $\frac{1}{10}$־יוצר פסאודו רנדומי, כאשר $n=O\left(\log m\right)$.
בשלב הראשון נרצה לנפח ל $x\in U_n$ $y\in\{0,1\}^{n+1}$ כך ש y יראה אקראי לחלוטין, אבל הניפוח של x ל y יהיה דטרמיניסטי.
בהנתן f כך ש $H_{avg}\left(f\right)\left(n\right)>s\left(n\right)$ נבנה $\varepsilon-PRG$ שהוא $g:\{0,1\}^n\to\{0,1\}^{n+1}$ עבור מעגלים בגודל קצת יותר קטן מ s. (נגדיר $\varepsilon=\frac{1}{s(n)}$)
נגדיר:

$$g\left(x_1,...,x_n\right)=\left(x_1,...,x_n,f\left(x\right)\right)$$

מדוע זה אכן יוצר כנ"ל? כי אם באמצעות מעגל קטן מ s ניתן להבדיל בין התפלגויות - נוכל באמצעות אותו מעגל לחשב את f...
ביתר פירוט:

$$P_1=Pr\left[C\left(x,f\left(x\right)\right)=1\right]$$

$$P_2=Pr\left[C\left(x,\underbrace{x_{n+1}}_{random\,bit}\right)=1\right]$$

נניח:

$$P1-P_2>\varepsilon$$

ונגדיר את C' שמקבל את $x_1,...,x_n$ ביטים ועוד ביט רנדומי x_{n+1}.
נריץ את $C\left(x_1,...,x_n,x_{n+1}\right)$, אם מקבלים את התשובה 1 - אז C בעצם אומר לנו ש"לדעתו" הביט שהוספנו הוא $f\left(x\right)$, ולכן נחזיר אותו, ואחרת נחזיר את $\overline{x_{n+1}}$.

10 שיעור 10

ראינו בעבר פיענוח לוקלי בעזרת מספר פולי־לוגריתמי של גישות.
לקוד WH (וולש הדמר) ראינו פיענוח לוקאלי בעזרת 2 שאילתות (אורך הקוד אקספוננציאלי).

שאלה פתוחה: האם יש קוד $f:\{0,1\}^n\to\Sigma^m$ עם קצב פולינומי (כאן פולינומי הכוונה גם למספרים שליליים במעריך הפולינום) אשר יש לו פיענוח בעזרת מספר קבוע של גישות? יודעים שעבור 2 גישות בלבד התשובה היא לא, כלומר אורך הקוד חייב להיות אקספוננציאלי.
עבור k גישות - בעצם ידועים קודים תת־מעריכיים (Sub-exponential)

תזכורת: $f:\{0,1\}^*\to\{0,1\}$ מקיימת $H_{avg}\left(f\right)\left(n\right)>s\left(n\right)$ אם אין מעגל בגודל קטן מ $s\left(n\right)$שמחשב את f על $\frac{1}{2}+\frac{1}{s(n)}$ מהקלטים באורך n.
הוכחנו שאם $f\in ExpTime$ ו $H\left(f\right)\left(n\right)>s\left(n\right)$ אז קיימת $g\in ExpTime$ ו $c>0$ כך ש $H_{avg}\left(g\right)\left(m\right)\geq s\left(\frac{m}{c}\right)$
(כלומר שאם יש פונק' קשה במקרה הגרוע, יש פונק' שהיא קשה בממוצע)

הגדרה 10.1 התפלגות D על $\{0,1\}^n$ היא ε-פסאודו רנדומית (מעתה יסומן פ"ר) למעגלים בגודל s אם ... וכו'

הערת סימון: כאן אנו מסמנים ע"י U_k דגימה מתוך התפלגות יוניפורמית על k ביטים.

הגדרה 10.2 פונקציה $g:\{0,1\}^* \to \{0,1\}^*$ היא $PRG-s(l)$ (דהיינו יוצר פ"ר) אם היא חשיבה בזמן $2^{O(l)}$ על קלטים בגודל l, ובהנתן קלט כזה מייצרת מחרוזת באורך $s(l)$ כך ש $g(U_l)$ היא $\frac{1}{10}$ פ"ר עבור מעגלים בגודל $s(l)^3$.

דוגמת "צעצוע"

טענה 10.3 תהי $f \in ExpTIme$ המקיימת $H_{avg}(f)(n) > n^4$ אזי הפונקציה g המקיימת:

$$g(z) = g(z_1, z_2, ..., z_l) = (z, f(z))$$

היא $(s(l) = l+1)$- יוצר פ"ר.

הוכחה: נניח שיש מעגל c בגודל $(l+1)^3 \geq$ שמבדיל בין $g(U_l)$ ובין U_{l+1} (מבדיל עם פרמטר $\frac{1}{10}$)
תחת הנחה זו נבנה מעגל c' שמחשב את fעל יותר מאשר $\frac{1}{2}+\frac{1}{l^4}$ מהקלטים.
בהנתן הקלט (z, z_{l+1}) שהוא מחרוזת בת $l+1$ ביטים, נפעיל את c עליהם, אם $c(z, z_{l+1}) = 1$ נחזיר z_{l+1}, ואחרת נחזיר את $\overline{z_{l+1}}$
נטען: אם בוחרים את z, z_{l+1} באופן מקרי, אז $Pr[c'(z, z_{l+1}) = f(z)] \geq \frac{1}{2}+\frac{1}{10}$.
מספיק להוכיח את הטענה ואז קיימת בחירה של z_{l+1} (נגיד 1) שעבורה ההסתברות ש c' מחשב נכונה את f גדולה מ $\frac{1}{2}+\frac{1}{10}$, וזה הרבה יותר גדול מ $\frac{1}{2}+\frac{1}{l^4}$, ו $l^4 \ll |c'|$. זאת בסתירה לקושי החישוב של f. **הוכחה:** (טענת העזר)
נגדיר:

$$p_g = Pr[c(g(U_l)) = 1]$$

כלומר ההסתברות שכשנותנים ל c את $(z, f(z))$ אז הוא יתן 1.
נגדיר גם:

$$p_u = Pr[c(U_{l+1}) = 1]$$

וזו ההסתברות ש c ישיב 1 כאשר נותנים לו z, u_1, כלומר בעצם $l+1$ ביטים שהוגרלו באקראי.
אנחנו יודעים מהנחתנו ש $p_g - p_u \geq \frac{1}{10}$.
נחשב:

$$Pr_{(z,z_{l+1})}[c'(z, z_{l+1}) = f(z)] = \frac{1}{2}\underbrace{Pr[c'(z, z_{l+1}) = f(z) \,|\, z_{l+1} = f(z)]}_{=p_g} + \frac{1}{2}\underbrace{Pr[c'(z, z_{l+1}) = f(z) \,|\, z_{l+1} \neq f(z)]}_{=Pr\left[c\left(z,\overline{f(z)}\right)=0\right]}$$

$$= \frac{1}{2}p_g + \frac{1}{2}Pr\left[c\left(z, \overline{f(z)}\right) = 0\right]$$
$$= \frac{1}{2}p_g + \frac{1}{2}\left(1 - Pr\left[c\left(z, \overline{f(z)}\right) = 1\right]\right)$$
$$= \frac{1}{2}p_g + \frac{1}{2} - \underbrace{\frac{1}{2}Pr\left[c\left(z, \overline{f(z)}\right) = 1\right]}_{A} - \underbrace{\frac{1}{2}Pr[c(z, f(z)) = 1]}_{B} + \underbrace{\frac{1}{2}Pr[c(z, f(z)) = 1]}_{=\frac{1}{2}p_g}$$

נבחין כי $A + B = p_u$ ולכן:

$$= p_g + \frac{1}{2} - p_u \geq \frac{1}{2} + \frac{1}{10}$$

■

וכיוון שראינו שבהנתן טענה זו מוכח הדרוש - סיימנו.

■

ניפוח של יוצר פ"ר

איך נייצר מספר גדול של ביטים שנדמה אקראי, עכשיו כשיש בידנו דרך לייצר אחד?
הרעיון הלא נכון - מ l הביטים הראשונים ייצרנו את הביט ה $l+1$.
אחר כך נפעיל את f על הביטים $(z_2,...,z_l,f(z))$ כדי לקבל את הביט ה $z+1$... וכך הלאה.
זה עובד עד גבול מסויים, כלומר עד מספר פולינומי של חזרות.
הרעיון הכן נכון - נתחיל מ $l=n^4$ ביטים, נמצא תת קבוצה שלהם I_1 שגדלה הוא n, ונכניס רק את הביטים הללו לתוך f, נכנה את התוצאה g_1. כעת נבחר תת קבוצה אחרת I_2 ונחשב באותו אופן את g_2. נעשה זאת $s(l)$ פעמים וכך נחשב את $g(z)=g_1,g_2,...,g_{s(l)}$.
כדאי לנו שהקבוצות לא יהיו דומות מדי, שכן זה יגרום לכך שהתוצאה תידמה אקראית יותר, ולכן נדרוש גם $\forall j\neq k\ |I_j\cap I_k|<d$ עבור d כלשהו.

הגדרה 10.4 בהנתן אוסף $\mathcal{I}$ של תת קבוצות של $[l]$ בגודל n כ"א, ובהנתן $f:\{0,1\}^n\to\{0,1\}$ נגדיר את הגנרטור של נועם ניסן ואבי ויגדרזון כך:

$$NW_f^{\mathcal{I}}(z)=\left(f\left(z_{I_1}\right),...,f\left(z_{|I_m}\right)\right)$$

כאשר $m=|\mathcal{I}|$.

.

הגדרה 10.5 נגדיר $-\ design$ $\left(\underbrace{l}_{length},\ \underbrace{n}_{bits\,per\,set},\ \underbrace{d}_{max\,intersection\,size}\right)$ אוסף של תת קבוצות של $[l]$, שכל אחת בגודל n והחיתוך בין כל שתיים אינו עולה על d.

טענה 10.6 (קיום תבניות) קיים אלגוריתם A אשר בהנתן (l,n,d) המקיימים ש $n>d$, $l>\frac{10n^2}{d}$, האלגוריתם רץ בזמן $2^{O(l)}$ ופולט תבנית (l,n,d) שגדלה הוא לפחות $2^{\frac{d}{10}}$ (למעשה זה בדיוק גדלה, כי בשלב זה נעצור את האלגוריתם)

הוכחה: משיקולים הסתברותיים - בבחירת שתי קבוצות אקראיות בגודל n מתוך l, תוחלת החיתוך תהיה $\frac{n^2}{l}$.
כעת נבחין כי אם תוחלת החיתוך היא ε אז בבחירת m קבוצות, עבור הקבוצה השניה תוחלת החיתוך תהיה ε, עבור השלישית לכל היותר 2ε וכך הלאה, ובסופו של דבר אם נבחר $\varepsilon=2^{-d}$ אז נידרש ל $m=2^{\frac{d}{2}}$ חזרות כדי לקבל רשימה מוצלחת (כלומר של קבוצות עם מעט חיתוכים).
מכיוון שחסם זמן הריצה שלנו מעריכי, אפשר לעבור פשוט על רשימת כל תתי הקבוצות ולבחור את המתאימות, השיקול ההסתברותי מבטיח לנו קבוצה מהגודל הדרוש וזמן הריצה מאפשר לנו לעבור על כולן ולכן זה מסיים את הוכחת הטענה. ■

טענה 10.7 (NW עובדים): תהי $\mathcal{I}$ תבנית (l,n,d) כך ש $|\mathcal{I}|=2^{\frac{d}{10}}$ ו $f:\{0,1\}^n\to\{0,1\}$ מקיימת:

$$H_{avg}(f)(n)>2^{2d}$$

אזי $NW_f^{\mathcal{I}}(U_l)$ הוא $\frac{1}{10}$-יוצר פ"ר עבור מעגלים בגודל $\frac{H_{avg}(f)}{10}$.

מסקנה 10.8 אם $H_{avg}(f)(n)>2^{\frac{n}{10}}$ וניקח $d=\frac{n}{20}$ ו $l=200n$ (לינאריים ב n) ונניח $f\in ExpTime$. אזי בהנתן הטענה, בזמן אקספ' ב l (וזה אקספ' ב n) אנחנו מחשבים את התבנית, לכל אחד מפלטים אנחנו מחשבים את f. ולכן סה"כ בזמן אקספ' ב n נקבל $2^{\frac{n}{200}}$ביטים שהם $\frac{1}{10}$-פ"ר למעגלים בגודל $\frac{2^{\frac{n}{10}}}{10}$.
כעת נניח שיש לנו מעגל רנדומי בגודל m^7 שמקבל קלט בגודל m ומחשב איזושהי פונקציה. נבחר את $n=700\cdot\log m$. כעת בזמן אקספ' ב n, ולכן פולינומי ב m, נוכל לייצר ביטים שנראים רנדומיים למעגל הזה. ואז נריץ את הגנרטור על כל המילים באורך n, זה מספר פולינומי ב m, נחשב על כל המילים הללו את הפלט של המעגל ונכריע לפי בחירת הרוב.

מה קיבלנו? שאפשר לבצע דה־רנדומיזציה של המעגל באמצעות היוצר הפ"ר הנ"ל! כלומר:

$$BPP \subseteq P$$

...

מסקנה 10.9 תהי $f \in ExpTime$

1. אם $H(f)(n) > 2^{\varepsilon n}$ אז $BPP = P$
2. אם $H(f)(n) > 2^{n^{\varepsilon}}$ אז $BPP \subseteq Quazi - P$ (קוואזי P זה n בחזקת משהו שהוא פולי לוגריתמי ב n)
3. אם $H(f)(n) > n^{\omega(1)}$ (אומגה כאן היא כל פונקציה ששואפת לאינסוף) אזי $BPP \subseteq Sub - Exp = \bigcap_{\varepsilon>0} TIME\left(2^{n^{\varepsilon}}\right)$

הוכחה: נניח בשלילה שקיים מעגל c כך ש $|c| \leq \frac{H_{avg}(f)(n)}{10}$ שמבדיל ב"הצלחה" $\leq \frac{1}{10}$ בין $U_m \sim z = (z_1, ..., z_m)$ (כאשר $m = |\mathcal{I}|$) ובין $NW_f^{\mathcal{I}}(U_l)$.
נסמן $NW_f^{\mathcal{I}}(U_l) = g = (g_1, ..., g_m)$
וכן $s = H_{avg}(f)(n)$
נבנה סדרה של משתנים מקריים:

$$\begin{aligned} z &= (z_1, z_2, z_3..., z_m) \\ &(g_1, z_2, z_3..., z_m) \\ &(g_1, g_2, z_3, ..., z_m) \\ &\vdots \\ &(g_1, ..., g_m) = g \end{aligned}$$

ולכל אחד מהם נסמן את ההסתברות ש c נותן עליו 1, כלומר:

$$\begin{aligned} p_0 &= Pr[c(z) = 1] \\ p_1 &= Pr[c(g_1, z_2, ..., z_m) = 1] \\ &\vdots \\ p_m &= Pr[c(g) = 1] \end{aligned}$$

והנחנו בשלילה כי:

$$p_m - p_0 \geq \frac{1}{10}$$

ולכן קיים i כך ש:

$$p_i - p_{i-1} \geq \frac{1}{10m}$$

טענה: מכאן נובע שיש מעגל c' בגודל $\geq \frac{s}{2}$ כך ש:

$$Pr[c'(g_1, ..., g_{i-1}) = g_i] \geq \frac{1}{2} + \frac{1}{10m} = \frac{1}{2} + \frac{1}{10 \cdot 2^d}$$

הוכחה: בדומה לדוגמת הצעצוע.
נרצה לבנות מ c' מעגל בגודל $\geq s$, המחשב את f על יותר מ $\frac{1}{2} + \frac{1}{10m}$ מהקלטים.

נתבונן באירוע $c'(g_1,...,g_{i-1})=g_i$, זהו בעצם:

$$c'\left(f\left(z_{|I_1}\right),...,f\left(z_{|I_{i-1}}\right)\right)=f\left(z_{|I_i}\right)$$

וזה מתקיים על לפחות $\frac{1}{2}+\frac{1}{10m}$ מה zים.
בהנתן l־יה של ביטים שמסומנת z, נקבע את z_1 להיות הביטים במקומות המתאימים ל I_i ו z_0 להיות שאר הביטים.
בה"כ I_i הם בדיוק n הביטים האחרונים ואז $z=z_0,z_1$.
אז מה שקיבלנו מטענה הקודמת ניתן לרישום כך:

$$Pr_{z_0,z_1}\left[c'\left(f\left(z_0,z_1\right)_{|I_1},...f\left(z_0,z_1\right)_{|I_{i-1}}\right)=f\left(z_1\right)\right]\geq\frac{1}{2}+\frac{1}{10m}$$

ולכן אפשר לקבוע z_0 ולתת להסתברות לרוץ על z_1 ולקבל עדיין:

$$Pr_{z_1}\left[c'\left(f\left(z_0,z_1\right)_{|I_1},...f\left(z_0,z_1\right)_{|I_{i-1}}\right)=f\left(z_1\right)\right]\geq\frac{1}{2}+\frac{1}{10m}$$

אנחנו יודעים שלכל $j<i$ אפשר לחשב את $f\left(z_0,z_1\right)_{|I_j}$ בזמן 2^d, מדוע? כי z_0 קבוע, ו I_j מכיל לכל היותר d ביטים מתוך z_1, ובעצם $f_{|I_j}$ היא פונקציה של d ביטים ביחס לכך.
נגדיר $c''(z_1)=c'\left(f\left(z_0,z_1\right)_{|I_1},...f\left(z_0,z_1\right)_{|I_m}\right)$ מעגל עם n קלטים המקיים $c''(z)=f(z)$ בסיכוי $\frac{1}{2}+\frac{1}{10m}$.
כעת מהו הגודל של c''? יש לנו $i-1$ מעגלים (בפרט פחות מ m) שכל אחד קטן מ 2^d, ועוד מעגל כללי c' שמקבל את פלטי המעגלים הקודמים, והוא בגודל $\frac{s}{2}$ לכל היותר.
זה מסתכם ל $\frac{s}{2}\approx\frac{s}{2}+2^{d+\frac{1}{10}d}\cdot 2^d\leq\frac{s}{2}+\underbrace{m}_{=2^{\frac{d}{10}}}$ וזה קטן מ s, בסתירה לקושי של f. ■

11 שיעור 11

(שיעור זה יסתמך בחלקו על הספר של ארורה וברק וכן על סיכומי ההרצאות של לוקה טרוויסאן)
נוכיח ש $Parity\notin AC_0$

קצת הגדרות פונקציית הזוגיות Parity מוכרת ופשוטה, ניתן להגדירה גם אם בוחרים $\{-1,1\}^n$ על ידי כפל כל הקואורדינטות זו בזו.
אם C מעגל בוליאני, עומקו $depth(C)$ מוגדר להיות המרחק הגדול ביותר בגרף המכוון המתאים מקלט לפלט.
AC_0 = מחלקת הפונקציות על n ביטים שניתן לחשב על ידי מעגל בעומק d מעומק $O(1)$ וגודל $Poly(n)\geq$
(כמובן ה $fan\,in$ אינו חסום, כי אחרת כל פלט לא יהיה תלוי ביותר ממספר קבוע של שערים).
המשפט שנרצה להוכיח אומר במילים: לכל מעגל בעופק חסום המחשב את פונקציות הזוגיות יש גודל סופר פולינומי.
שאלה: ל dנתון, מהו הגודל המזערי של מעגל המחשב את $parity$?
תשובה: להלן בניה מגודל $n\cdot 2^{n^{\frac{1}{d-1}}}$
על ידי פעולות פשוטות ותוך הגדלה של המעגל בגורם קבוע לכל היותר, ניתן להניח שהמעגלים שבהם אנו עוסקים הם מובנים כדלקמן:

1. אין שערי שלילה, במקום זאת יש $2n$ קלטים שהם $x_1,\neg x_1,...,x_n,\neg x_n$ (על ידי שימוש בחוקי דה מורגן אפשר לחלחל את כל שערי השלילה עד לשערי הקלט)

2. המעגל מסודר ברמות מתחלפות. בכל רמה כל השערים הם שערי $\vee$ ולחילופין שערי $\wedge$ (אפשר להשיג זאת בקלות, תרגיל פשוט)

3. מעגל בעומק 2 שברמה התחתונה יש $\vee$ וברמה שמעליו יש $\wedge$מחשב בדיוק פונקציה בוליאנית מצורת CNF ואם $\wedge$ ואז $\vee$אז מצורת DNF. איך נראית פונקצית זוגיות בעומק 2? ניקח איווי על כל טבלת האמת במקומות שבהם הזוגיות היא 0 - פסוקית לכל ביטוי, והפסוקית בודקת האם בדיוק התקבל הביטוי המתאים. זה יהיה מצורת DNF ולכן ניתן למימוש בעומק 2. כאן יהיו 2^{n-1} גורמים בביטוי ה DNF.

רעיון הבניה

נתחיל בעץ מעומק d ודרגת כניסה של $n^{\frac{1}{d}}$, שבכל קודקוד פנימי שלו מופיע שער $parity$. זהו בעצם שימוש בקומוטטיביות ואסוצ' של חיבור מודולו 2, במקום לחבר את כל הביטים במכה בשער זוגיות אחד, אנחנו מפצלים לקבוצות קטנות, שגודל כ"א $n^{\frac{1}{d}}$, ומחשבים זוגיות עבורן וכך הלאה.

נחליף עתה כ"א משערי הזוגיות שבהם השתמשנו במימוש DNF/CNF כלשהו שלו, זהו מעגל בעומק 2 ובגודל $2^{n^{\frac{1}{d}}}$, כעת המעגל החדש בעומק $2d$, אבל אם נבחר לסירוגין DNF ו CNF, נקבל רמות עוקבות של $\vee$ ו $\wedge$ שנוכל לצמצם ולקבל עומק d.

משפט 11.1 לכל מעגל בוליאני עם n קלטים ועומק dהמחשב את פונ' הזוגיות יש גודל $\Omega\left(2^{n^{\frac{1}{4d}}}\right)$.

ההוכחה - אנליזה הרמונית על הקוביה הדיסקרטית.

לכל פונקציה $f:\{-1,1\}^n\to\mathbb{R}$ יש הצגה אחת ויחידה כפולינום מולטיליניארי ב $x_1,...,x_n$.

(פולינום מולטיליניארי הוא כזה שבו כל משתנה מופיע ממעלה 0 או 1 ואז אפשר לכתוב זאת כך $f=\sum_S a_S\prod_{i\in S}x_i$)

סקיצה של "הוכחה": מביטים באוסף הפונקציות $\{f:\{-1,1\}^n\to\mathbb{R}\}$ כמרחב וקטורי ממשי.

המרחב הוא 2^n מימדי, ומגדיירים את המכפלה הפנימית כך $\langle f,g\rangle=\frac{1}{2^n}\sum_x f(x)\,g(x)$

להלן בסיס אורתונורמלי שלו:

הפונקציות $\prod_{i\in S}x_i$ ע"פ $S\subseteq[n]$ מהוות בסיס אורת' למרחב זה.

מכיוון שכך, אם ניקח הצגה של פולינום מולט' $f=\sum_S a_S\prod_{i\in S}x_i$ אז נוכל לחשב את מקדמי פורייה $\hat{f}(S)=g,\langle f,\prod_{i\in S}x_i\rangle$

נסמן $\prod_{i\in S}x_i=\chi_S$

תחת הסימונים לעיל מקבלים את הצגת הפולינום כ $\sum_{S\subseteq[n]}\hat{f}(S)\,\chi_S$

ונוכל לייחס לכל פונקציה ממשית על הקוביה - דרגה, אותה נגדיר כך $deg(f)=\max_{\hat{f}(S)\neq 0}|S|$

למה 11.2 יהי C מעגל בוליאני מעומק d וגודל s. אז יש $g:\{-1,1\}^n\to\mathbb{R}$ מדרגה $\geq O\left((\log s)^{2d}\right)$ כך ש g,C מתלכדים על לפחות $\frac{3}{4}$ מהקלטים.

.

למה 11.3 תהי $g:\{-1,1\}^n\to\mathbb{R}$ המתלכדת עם פונק' הזוגיות על לפחות $\frac{3}{4}$ מהקלטים, אזי $deg(g)\geq\Omega(\sqrt{n})$

איך נובע המשפט מהלמות? **הוכחה:** (המשפט בהנחת הלמות) נניח C מעגל בעומק d המחשב זוגיות, יהי s גודלו של C, צ"ל $s\geq\Omega\left(2^{\frac{1}{4d}}\right)$.

בעזרת הלמה הראשונה נמצא פונ' $g:\{-1,1\}^n\to\mathbb{R}$ מדרגה $\geq O\left((\log s)^{2d}\right)$ המתלכדת עם C על יותר מ $\frac{3}{4}$ מהקלטים.

כלומר $Pr[g(x)=\prod x_i]\leq\frac{3}{4}$ ולכן לפי הלמה השניה נקבל $deg(g)\geq\Omega(\sqrt{n})$ ומהעברת אגפים

$$(\log s)^{2d}\gg deg(g)\gg n^{\frac{1}{2}}$$

כעת מחילוץ s מקבלים את הדרוש. ■

. **הוכחה:** (למה 1) הערה - מעתה נרשום כאילו הפונ' הבוליאניות מקבלות $1,0$ ולא $-1,1$.

הרעיון: להחליף כל שער של $\vee$או $\wedge$ בפולינום ממעלה נמוכה ש"בדרך כלל" יקבל רק ערכי $0,1$ וכמובן את הערך הנכון.

לדוגמה - אם יש $\bigvee_{i=1}^{k} x_i$ אז נחליף אותו ב $1-\prod_{i=1}^{k}(1-x_i)$, הבעיה היא שלא חסמנו כך את הדרגה, אבל מאידך לא ניצלנו את החופש לקבל ערכים ממשיים, וגם לעיתים לשגות.

טענת עזר: לכל k, ε יש התפלגות G על פולינומים ממשיים $\{0,1\}^k \to \mathbb{R}$כך ש:

1. לכל פולינום בטווח של G יש דרגה $O\left(\log\frac{1}{\varepsilon}\cdot\log k\right)$ ולכל $x\in\{0,1\}^k$ מתקיים:

$$Pr_{g\sim G}\left[g\left(x_1,...,x_k\right)=\bigvee_{i=1}^{k}x\right]\geq 1-\varepsilon$$

הוכחת טענת העזר: ההגרלה תיראה כך: אנו נבחר באקראי משפחה F של תת קבוצות של $[k]$ והפולינום המתאים יהיה:

$$g\left(x_1,...,x_k\right)=1-\prod_{S\in F}\left(1-\sum_{i\in S}x_i\right)$$

אם כל $x_i=0$ אז $g\left(0,0....\right)=0$ וזה בסדר.

אם יש איזשהו $x_j=1$ אז נצליח $\Leftarrow$ יש $S\in F$ כך ש $\#\left\{x_j : j\in S,\ x_j=1\right\}=1$ (שכן רק במקרה כזה יהיה 0 במכפלה ונקבל 1 בסך הכל)

חשבון עזר: יש שתי קב' מקריות מגודל a,b חלקיות ל $[m]$, מהי ההסתברות שחיתוכן מכיל בדיוק איבר 1?

נניח ש b קבועה, אזי השאלה בכמה מבין $\binom{m}{a}$ הבחירות האפשריות גודל החיתוך הוא בדיוק 1? כלומר $=\frac{b\binom{m-b}{a-1}}{\binom{m}{a}}$

$$b\frac{a}{m-a-b+1}\frac{\binom{m-b}{a}}{\binom{m}{a}}$$

$$=\frac{ab}{m-a-b+1}\cdot\frac{(m-b)(m-b-1)...(m-b-a+1)}{m\cdot(m-1)...(m-a+1)}$$

נניח $a.b\ll m$ ואז מה שקיבלנו הוא:

$$\approx(1-o(1))\frac{ab}{m}\left(1-\frac{b}{m}\right)^a\approx(1-o(1))\frac{ab}{m}e^{\frac{-ab}{m}}$$

ולכן אם $\frac{1}{2}<\frac{ab}{m}<1$ אז ההסתברות לעיל גדולה $\frac{1}{10}$ (בערך).

ולכן נבחר את F באופן הבא: נבחר באקראי $\log\frac{1}{\varepsilon}$ קבוצות מגודל 1, וכן $\log\frac{1}{\varepsilon}$ קבוצות מגודל 2, ... מגודל 4 ... מגודל 8 וכך הלאה, נקבל $\log k\log\frac{1}{\varepsilon}$ קבוצות.

בהנתן קבוצה מגודל l נתבונן בקבוצות מגודל $t=2^j$ כך ש $\frac{k}{l}\geq t\geq\frac{k}{2l}$ ואז יתקיים התנאי מחשבון העזר ולכן ההסתברות להצלחה תהיה גדולה מעשירית, וההסתברות לכישלון קטנה מ 0.9. נבדוק את כל $\log\frac{1}{\varepsilon}$ הקבוצות הרלבנטיות, ואז ההסתברות לכישלון בכולן יחד תהיה $\varepsilon=0.9^{\log\frac{1}{3}}$, הפולינום הוא מכפלה של פולינומים לינאריים המתאימים לקבוצות ב F. יש $\log k\log\frac{1}{\varepsilon}$ קבוצות, ולכן זו מעלת הפולינום המקסימלית, וזה מוכיח את טענת העזר.

מכאן נובעת למה 1:

נשתמש בטענת העזר לכל שער ב C עם $k=s$ ו $\varepsilon=\frac{1}{4s}$ ואז נקבל $deg\left(h\right)\leq O\left(\log^{2d}s\right)$ (כאן h הוא הפולינום שבנינו)

מדוע ההסתברות לשגיאה קטנה מרבע? כי כדי שזה יקרה יש בהכרח לפחות שער אחד שבו ה g המקורי היא שונה מ $\vee$, אבל זה קורה בהסתברות קטנה מ$\varepsilon=\frac{1}{4s}$ ועל ידי חסם האיחוד, ההסתברות הכולל לשגיאה קטנה מ $\frac{1}{4}$. ■

. **הוכחה:** (למה 2)

תהי $A\subseteq\{-1,1\}^n$ קבוצת הוקטורים x המקיימים $g\left(x\right)=\prod x_i$ (כלומר שעליהם g מסכימה עם פונק' הזוגיות). נביט במרחב הוקטורי V של כל ההעתקות $f:A\to\mathbb{R}$. זהו מרחב וקטורי במימד $|A|$.

<u>שתי אבחנות</u>

1. כל $f\in V$ ניתן להציג $f\left(x\right)=\sum_{S\in[n]}\alpha_S\chi_S$ לכל $x\in A$.
2. אם $x\in A$ אז $g\left(x\right)=\prod x_i$ (בהגדרה)

כעת אם $x\in A$ ו $S\subseteq[n]$ מה ניתן לומר על $g\left(x\right)\prod_{i\in s}x_i$? שזה בדיוק $\prod_{j\notin S}x_i$... (מתקבל מהעברת אגפים)

לכן אם $f \in V$ אז ניתן לרשום אותו (לכל $x \in A$) ע"י

$$f(x) = \sum \alpha_S \chi_S = \sum_{|S| \leq T} \alpha_S \chi_S + g(x) \sum_{|S| > T} \alpha_S \chi_{\overline{S}}$$

נאמר שמעלת g היא t ונסמן את המשלימות של ה S ע"י הקבוצות Z ונוכל לרשום:

$$\sum_{|S| \leq T} \alpha_S \chi_S + \underbrace{g(x)}_{deg=t} \sum_{|Z| > n-t} \alpha_Z \chi_Z$$

מכך יתקבל שדרגת f היא קטנה או שווה ל $\max\{T, n-T+t\}$ ואם נבחר $T = \frac{n+t}{2}$ נקבל כי $deg(f) \leq \frac{n+t}{2}$.
אבל זה מוזר, כי מצד אחד Aהיא קבוצה גדולה מאד, בפרט גודלה הוא $\frac{3}{4}2^n$, וזה מימד V.
מצד שני - ראינו שניתן לבטא כל $f \in V$ ע"י ביטוי שדרגתו לכל היותר $\frac{n+t}{2}$, והמימד של מרחב הפולינומים מדרגה זו הוא $\sum_{k=0}^{\frac{n+t}{2}} \binom{n}{k}$ ולכן קיבלנו:

$$\frac{3}{4}2^n \leq \sum_{k=0}^{\frac{n+t}{2}} \binom{n}{k}$$

אבל כדי שזה יתקיים חייב להיות שבביטוי הימני עברנו את הביטוי הבינומי האמצעי ב $\Omega(\sqrt{n})$ (משחקים עם האלגברה וזה יוצא).
וזה מסיים את הוכחת למה 2. ■

הרעיון העיקרי בהוכחתו של האסטאד (הוכחה אלטרנטיבית לאותה בעיה) היא הגבלה מקרית של פונקציה בוליאנית (Random restriction).
בהנתן:

$$f : \{0,1\}^n \to \mathbb{R}$$

הגבלה מקרית שלה היא התפלגות על פונקציות על פחות משתנים שמתוארת כך:
נבחר פרמטר $0 < \rho < \frac{1}{2}$ (זה חלק מההגדרה) לכל $1 \leq i \leq k$ בהסתברות ρ ובאופן ב"ת נציב ל x_i את 0, בהסתברות ρ נציב את 1, ובהסתברות $1-2\rho$ משאירים את x_i כמשתנה.
קיבלנו פונקציה חדשה שיש בה באופן כללי פחות משתנים, וחלקה נקבע באופן מקרי.
הלמה הבסיסית של האסטאד אומרת שאם f מתוארת ע"י נוסחה ב $t-CNF$ ואם מבצעים עליה ρ-הגבלה מקרית, אזי בהסתברות גבוהה ניתן לרשום את הפונקציה המצומצמת ב $s-DNF$ לא גדולה.