שיטות הצפנה

מבוסס על הרצאותיו של פרופ' מיכאל בן אור

סימסטר ב', תש"ע

הערה 0.1 סוכם ע"י אלון גונן (בעזרתו של שיר פלד). תיקונים יתקבלו בברכה בכתובת $alongnn@gmail.com$.

הערה 0.2 דרישות הקורס: ינתנו תרגילים, חלקם להגשה אך רובם לא. בחינה בכתב (קלה לדברי בן־אור). ציון ינתן על בסיס התרגילים והבחינה.

1 רקע היסטורי

הצורך להצפין קיים כבר שנים רבות. $Ceasar$ יצר מעגל שמגדיר סיבוב של הא"ב וכל שצריך הוא קונבנציה בין השולח למקבל היודעים את גודל הסיבוב. לא קשה מדי לשבור שיטה שכזו ע"י נסיון חוזר של סיבובים שונים עד שמקבלים טקסט בעל משמעות. שיטה שונה היא להגדיר פרמוטציה על האתיות בא"ב המקורי. בהסתמך על העובדה שקייימות 26 אותיות בשפה, קיימות $26!$ אפשרויות שונות. זה כבר מספר גדול יותר אך זה עדיין ניתן לשבירה. קיימות ווריאציות נוספות אשר כולן ניתנות לשבירה.

אחרי מלחמת העולם הראשונה המחקר בתחום ההצפנה בעיקר ע"י גופים צבאיים התפתח ללא הפסקה. מחקרים אלו לא התפרסמו. ב$Helman$ 1976 פיתח מערכת הצפנה משוכללת. אחריה בשנת 1977 פורסמה שיטת ההצפנה הידועה RSA. ה־NSA ניסה לעכב את פרסומם של הדברים הללו אבל זה היה מאוחר מדי. עד אמצע שנות ה־90 ארה"ב אסרה על הפצה של תוכנות קריפטוגרפיות. לאחר מכן כבר התחיל מחקר גלוי.

הקורס יסקור את השיטות החל מסוף שנות ה־70. הקורס לא יעסוק בשאלה האם פרוטוקול מסויים הוא בטוח ברמת המימוש. קיים פער עדין בין הרעיון של הפרוטוקול לבין מימושו. אנו נדון ברעיונות ולא נעסוק במימוש.

2 מהי הצפנה בטוחה?

2.1 הצפנה בטוחה, שיטת Shannon

שנון הוא הראשון שהמציא את תורת האינפורמציה. ב1947 פרסם מאמר בו הגדיר מהי שיטת הצפנה בטוחה לחלוטין - Information Theoretic Security.

נתאר את מבנה הבעיה. כמו תמיד, יש לנו את Bob, Alice, Eve (נסמן ב־B, A, E). אליס רוצה לשלוח הודעה לבוב. איב מעוניינת להאזין. ב.ה.כ ההודעות הן בביטים. נסמן ב־$m \in \{0,1\}^n$ את ההודעה שאליס מעוניינת לשלוח. קל להראות שאם בוב ואליס לא נפגשו מעולם ולא סיכמו בינהם קונבנציה כלשהי, לא קיימת דרך כך שאליס תעביר לבוב את ההודעה לבוב מבלי שאיב יוכל לפענח אותה (לא נוכיח זאת אך נסתפק בהבנה לפיה לבוב אין כל אינפוררמציה שלא קיימת ברשות איב).

נתאר כעת את שיטת שנון. אנו נניח **הנחה** לפיה לבוב ואליס קיים מפתח סודי משותף $k \in \{0,1\}^n$. כמו כן נניח כי k נבחר בהתפלגות אחידה על התחום $\{0,1\}^n$. את ההתפלגות האחידה על $\{0,1\}^n$נסמן ב־U_n. אם כך, בוב ואליס יודעים את k. איב אינה יודעת את k אך יודעת ש־k מתפלג לפי U_n. שנון הציע שבהנתן הודעה m אליס תשלח את $c = m \oplus k$. בוב מצידו יקבל את c וכאמור יודע את k, יוכל לפענח את ההודעה המקורית m ע"י חישוב $m = c \oplus k = m \oplus k \oplus k = m$.

הערה 2.1 נזכיר שהפעולה xor אותה אנו מסמנים ב־$\oplus$ על זוג ביטים a, b שקולה לפעולה $(a+b)\%2$. בהינתן וקטורים של ביטים a, b פעולת xor עליהם מתבצעת ביט ביט.

לשיטה שתוארה לעיל קוראים $One\,time\,pad$. מיד נטען (ונוכיח) כי שיטת ההצפנה הזאת הינה בטוחה. לפני כן נגדיר מהי הצפנה בטוחה.

הגדרה 2.2 הצפנה בטוחה

תחילה נגדיר **הצפנה תקינה.** נגדיר העתקה E מהתחום שהינו מכפלה קרטזית של מפתחות (ווקטורים ב־R^r) והודעות מקוריות (ווקטורים ב־R^n) לטווח שמוגדר להיות הודעות נשלחות (ווקטורים ב־R^l). מנגד נגדיר העתקה D מהתחום שהינו מכפלה קרטזית של מפתחות והודעות נשלחות לטווח ־ הודעה מקורית. כלומר,

$$E : \{0,1\}^r * \{0,1\}^n \to \{0,1\}^l$$

$$D : \{0,1\}^r * \{0,1\}^l - > \{0,1\}^n$$

ונדרוש,

$$D_k(E_k(m)) = m$$

כלומר, אנו דורשים שמפענוח ההודעה המוצפנת מתקבלת ההודעה המקורית.

בהנתן מערכת (E, D) אנו נאמר שהמערכת **בטוחה** לחלוטין אם במשחק הבא הסתברות ההצלחה של איב היא לכל היותר $\frac{1}{2}$: איב בוחרת שתי הודעות $x_0, x_1 \in M = \{0,1\}^n$. ומוסרת אותן ל־$A$. A בוחרת ביט מקרי $b \in \{0,1\}$ ולפיו נקבע $C = E_{U_n}(x_b)$ ומוסרת את C ל־E. איב מנצחת אם היא מצליחה למצוא את b.

הערה 2.3 בסיכומים של בועז ברק מאוניברסיטת $princeton$ מופיעה תחילה ההגדרה הבאה עבור הצפנה בטוחה:

סכימת ההצפנה (D, E) (הכינוי סכימת הצפנה טומן הנחה על תקינותה) תקרא הצפנה בטוחה אם לכל שתי הודעות x, x' מתקיים $E_{U_n}(x) \equiv E_{U_n}(x')$ (כלומר ההתפלגויות המתארות את ההצפנות של כל שתי הודעות הינן זהות).

בהנתן ההגדרה הזאת, בועז מראה כי ההגדרה הנ"ל שקולה להגדרה באמצעות משחק.

טענה 2.4 שיטת שנון בטוחה.

לפני שנוכיח את הטענה נראה את הלמה הבאה.

למה 2.5 לכל $x \in \{0,1\}^n$ המשתנה המקרי $C = E_{U_n}(x)$ מתפלג באופן אחיד על $\{0,1\}^n$.

הוכחה: צ"ל כי לכל $C \in \{0,1\}^n$ מתקיים:
באופן שקול צ"ל שלכל $C \in \{0,1\}^n$:

$$Pr_{k \in U_n}[C = x \oplus k] = Pr_{k \in U_n}[C \oplus x = k] = \frac{1}{2^n}$$

וזה מתקיים מפני שאנו שואלים מה ההסתברות עבור k המתפלג לפי התפלגות אחידה, להיות זהה לווקטור שנקבע מראש (הווקטור $C \oplus k$). הסתברות זו שווה כמובן לפי הגדרה ל־$\frac{1}{2^n}$. ■

כעת, נחזור להוכחת הטענה: **הוכחה:** לפי הלמה C אינו תלוי ב־x_0, x_1 ולכן איב תוכל לכל היותר להסיק מתוך ההתפלגות האחידה על התחום $\{0,1\}$. מכאן סיכוייה לבחור את הביט הנכון בהתפלגות אחידה היא בדיוק $\frac{1}{2}$. באופן פורמאלי, נסמן ב־$i \in \{0,1\}$ את השערתה של איב (לפי נוסחת בייס):

$$Pr[x_i = x_b | C = c] = \frac{Pr[C = c, x_i = x_b]}{Pr[C = c]} = \frac{\frac{1}{2^n} * \frac{1}{2}}{\frac{1}{2^n}} = \frac{1}{2}$$

■

הערה 2.6 ברגע ש־C אינו תלוי ב־x יש לנו הצפנה בטוחה לחלוטין וכאן יש לנו מקרה כזה.

אז מה בעייתי בשיטה הזאת? המפתח חייב להיות **חד פעמי.** אם משתמשים פעמיים באותו מפתח. נסביר מדוע. אם אנו מצפינים שתי הודעות m_1,m_2 באמצעות אותו המפתח k נקבל $c_1 = m_1 \oplus k, c_2 = m_2 \oplus k$. במצב זה, $c_1 \oplus c_2 = m_1 \oplus m_2$, כלומר איב יכולה לפענח את שרשור ההודעות ומכאן בקלות את ההודעות עצמן. לכן, עבור 100 הודעות אנו נדרשים למפתח באורך $100n$. אם באופן כללי ניתן לדרוש פחות? כלומר, האם ניתן למצוא דרך בה אפשר לחסוך באורך המפתחות? מסתבר שלא.

טענה 2.7 ($Shannon$) בכל שיטה בטוחה לחלוטין אורך המפתח הינו לפחות אורך ההודעות המוצפנות.

הוכחה: נניח שאורך המפתח בגודל קטן או שווה $n-1$. לשם הנוחיות נתמקד במקרה שאורך המפתח שווה ל־$n-1$. נתבונן בהתפלגות $E_{U_{n-1}}(0^n)$. נסמן ב־S_0 את התומך של ההתפלגות הזו. מהנחותינו קיימים 2^{n-1}מפתחות ומכאן נובע $|S_o| \leq 2^{n-1}$. כעת, לכל מפתח קבוע $k \in U_{n-1}$ נתבונן באוסף ההצפנות $E_k(x)$ כאשר x עובר על $\{0,1\}^n$. ההעתקה חייבת להיות חח"ע כדי שנוכל לפענח (אם שתי הודעות תועקתנה לאותה הודעה מוצפנת לא נוכל לפענח). לכן, תמונת ההעתקה הזו הינה בגודל 2^n. לכן, עבור k הנ"ל קיים x_0 כך ש־$E_k(x_0) \notin S_0$. לכן, התומך של $E_{U_{n-1}}(x_0)$ איננו זהה לתומך של$E_{U_{n-1}}(0^n)$ כי יש נקודה ב־$E_{U_{n-1}}(x_0)$ אשר לא מופיעה ב־$E_{U_{n-1}}(0^n)$. מכאן התפלגויות אלו שונות.

נשים לב כי לפי ההגדרה המקורית עבור הצפנה בטוחה (לפי בועז ברק כפי שצויין לעיל) הוכחתנו הסתיימה מכיוון שהראינו כי קיימות שתי הודעות x, x_0 כך שההתפלגויות המתארות את ההצפנות שלהן אינן זהות.

נוכיח לפי גישת המשחק. מדוע זה סותר את הבטיחות לפי גישת המשחק? כי כעת נוכל לתת את x_0 ואת 0^n לאליס, ובהסתברות שאינה אפס־ נניח ϵ ־ אליס תשתמש ב k לעיל ותיתן לנו את $E_k(x_0)$, אשר ידוע לנו שאינו הצפנה של 0^n. בכל מקרה אחר נענה באופן אקראי, ולכן הסתברות ההצלחה בניחוש היא $\frac{1}{2} + \epsilon > \frac{1}{2}$ ומכאן שההצפנה אינה בטוחה־לחלוטין. ■

הגדרה 2.8 אם p, q התפלגויות על מרחב X, כלומר $p, q : X \to [0,1]$ ומתקיים $\sum_{x\in X} p(x) = \sum_{x\in X} q(x) = 1$, אזי נסמן את המרחק בין ההתפלגויות:

$$\Delta(p,q) = \max_{T\subseteq X} |p(T) - q(T)|$$

למה 2.9

$$\Delta(p,q) = \frac{1}{2}\sum_{x\in X} |p(x) - q(x)|$$

הוכחה: בכל מקום שבהם $p(x) \neq q(x)$ ־ אנחנו מחשבים את המרחק ביניהם. מכאן שסך הכל אנחנו מחשבים את סכום ההפרשים בנקודות שבהן $p > q$ ועוד סכום ההפרשים בנקודות שבהן $q > p$, כיוון שהם מסתכמים לאותו הדבר ־ הסכומים הללו שווים, ולכן כשנחלק את סכומם לשניים נקבל אחד מהם, וזהו בעצם $\Delta(p,q)$. ■

הערה 2.10 (ברק) באופן שקול, עבור התפלגויות X, Y על המרחב $\{0,1\}^n$, אנו נאמר שהמרחק בינהן הוא לכל היותר ϵאם מתקיים

$$|Pr[A(X) = 1] - Pr[A(Y) = 1]| \leq \epsilon$$

לכל $A : \{0,1\}^n \to \{0,1\}$.

הגדרה 2.11 נאמר שהמערכת היא ϵ בטוחה סטטיסטית, אם הסתברות ההצלחה (במשחק עם אליס שהגדרנו עבור הצפנה בטוחה) היא קטנה או שווה ל $\frac{1}{2} + \epsilon$. באופן שקול ניתן לדרוש שלכל זוג הודעות x,x', המרחק בין ההתפלגויות $E_{U_n}(x), E_{U_n}(x')$ הוא לכל היותר 2ϵ (ההגדרה השקולה מתוך הרצאותיו של ברק).

הערה 2.12 בהרצאותיו של בועז ברק מופיעה המוטיבציה הבאה להגדרה האחרונה: הראינו כי לא ניתן להצפין עם מתפחות באורך קטן מאורך ההודעות. כעת נרצה לבחון מחדש את ההנחות שמסתתרות מאחורי התוצאות אליהן הגענו ולמצוא האם אנו יכולים לבצע איזה שהיא $Relaxation$ של ההנחות כך שנוכל לשנות את התוצאות (כלומר להצפין עם מפתחות קצרים יותר). נשאלות שתי שאלות לגבי ההגדרה האחרונה. הראשונה - האם ההגדרה הנ"ל עדיין מבטיחה מערכת בטוחה? ובכן, התשובה היא כן. עבור ערכי ϵ מאד קטנים אנו מקבלים מערכת הצפנה בטוחה ליישומים רבים. השאלה השנייה היא האם ה-$Relaxation$ הנ"ל יאפשר לנו להפחית בדרישות לגבי אורך המפתח. אנו נראה בהמשך שלא.

כעת בהנתן התפלגויות p, q כך ש $\Delta(p,q) \neq 0$ מתבצע הניסוי הבא:
מטילים מטבע ובוחרים באחת מן ההתפלגויות, ודוגמים לפי ההתפלגות שנבחרה מספר, ומציגים לאיב את ה x שנבחר.
שואלים את איב מאיזו התפלגות נלקחה הדגימה, ונשאלת השאלה - מה הסתברות ההצלחה הטובה ביותר שאיב יכול לקוות לה.
מה הסתברות הצלחה?
נבהיר כי במסגרת הדיון שלנו ההתפלגויות p, q מתאימות להתפלגויות $E_{U_n}(m_0), E_{U_n}(m_1)$ עבור הודעות כלשהן m_0, m_1. מהמשפט הבא נקבל את מסקנתו של ברק אודות השקילות בהגדרה של מערכת בטוחה סטטיסטית.

טענה 2.13 אפשר לנצח בהסתברות $\frac{1}{2} + \frac{\epsilon}{2}$ אם"ם $\Delta(p,q) \geq \epsilon$

נניח תחילה כי $\Delta(p,q) \geq \epsilon$
האסטרטגיה הכי סבירה היא: בהנתן x, איב תשאל האם $p(x) \geq q(x)$ ואם כן - ננחש p, ואחרת ננחש q. נראה שאסטרגיה זו מבטיחה הצלחה בהסתברות הדרושה.
נגדיר $T = \{x \in X : p_x \geq q_x\}$
בהסתברות 0.5 נדגום לפי p. במצב זה איב תנחש נכונה, אם $x \in T$.
בהסתברות 0.5 נדגום לפי q. במצב זה איב תנחש נכונה אם $x \notin T$.
לכן, ההסתברות לתשובה נכונה תהיה

$$\frac{1}{2}p(T) + \frac{1}{2}(1 - q(T)) = \frac{1}{2} + \frac{1}{2}(p(T) - q(T)) = \frac{1}{2} + \frac{1}{2}\Delta(p,q) \geq \frac{1}{2} + \frac{\epsilon}{2}$$

ננסה להראות כעת את הכיוון השני - נניח שאנו יכולים לנצח בהסתברות $\frac{1}{2} + \frac{\epsilon}{2}$ ונראה ש $\epsilon \leq \Delta(p,q)$.
נסמן את ההתפלגויות ב $p_0 = p, p_1 = q$.
איב מפעילה אלגוריתם A כך שבהנתן $x \in X$ אז $A(x)$ יתן 0 או 1.
נגדיר

$$r_{i,j} = Pr\left(A_{x\, is\, sampled\, from\, p_i}(x) = j\right)$$

כלומר $\frac{r_{0,0}+r_{1,1}}{2}$ היא ההסתברות לניחוש נכון ו $\frac{r_{0,1}+r_{1,0}}{2}$ היא ההסתברות לשגיאה.
ברור ש $r_{0,0} + r_{0,1} = 1$, כיוון שהאלגוריתם תמיד יחזיר תשובה (נכונה או לא), ובאופן דומה $r_{1,0} + r_{1,1} = 1$.
ואז $\frac{r_{0,0}+r_{1,1}}{2} \geq \frac{1}{2} + \frac{\epsilon}{2}$ ומכאן $r_{0,0} + r_{1,1} \geq 1 + \epsilon$.
נחסר מכך את אחד השוויונות לעיל ונקבל:

$$r_{0,0} + r_{1,1} - r_{1,0} - r_{1,1} \geq \epsilon \Rightarrow r_{0,0} - r_{1,0} \geq \epsilon$$

נגדיר:

$$T = \{x \in X : A(x) = 0\}$$

ואז לפי הגדרה מתקיים:

$$p_0(T) = r_{0,0}$$
$$p_1(T) = r_{1,0}$$

ולכן עבור T לעיל מתקבל לפחות המרחק ϵ, ובפרט:

$$\Delta(p,q) \geq \epsilon$$

מסקנה 2.14 אם A הוא אלגוריתם ההבחנה הטוב ביותר בין ההתפלגויות p,q, כלומר נותן סיכוי הצלחה אופטימלי במשחק ההבחנה, אזי הסתברות ההצלחה של A היא בדיוק $\frac{1}{2}\Delta(p,q)$.

משפט 2.15 (אותו משפט של שנון, חוזרים עליו) בהצפנה בטוחה לחלוטין מרחב המפתחות גדול או שווה בגדלו למרחב ההודעות.

הוכחה: נניח כי אורך המפתחות הוא $n-1$. אזי $|k| = 2^{n-1}$.
בוחרים הודעה m_0 ומגדירים:

$$S = \{E_k(m_0) \,|\, k \in K\}$$

ברור כי $|S| \leq |K|$ולכן $|S| \leq 2^{n-1}$
בהנתן מפתח כלשהו k_0 מתקיים כפי שכבר טענו $|\{E_{k_0}(x) \,|\, x \in M\}| = 2^n$ (ההעתקה חייבת להיות חח"ע)
לכן, $\exists\, E_{k_0}(m_1) \notin S$ ולכן קיים m_1 כך ש $supp\{E_{k_0}(x) | x \in M\} \setminus S \neq \emptyset$
ההתפלגויות $E_K(m_0)$ ו $E_K(m_1)$ אינן שוות, ולכן יש מרחק שאינו אפס ביניהן, ולכן ההסתברות להצלחה המשחק אם ננקוט את האסטרטגיה הקודמת תהיה גדולה מחצי, בפרט - חצי ועוד מחצית המרחק ביניהן. ■

(חזרנו פה על ההוכחה הקודמת למשפט שנון במילים אחרות)
אם אורך מפתח ההצפנה $n-1 \geq$ להודעות באורך n, מתוך $\{0,1\}^n$, אז $|K| \leq \frac{|M|}{2}$.
אזי לכל $k_0 \in K$ לפחות חצי מ m מקיימים $E_{k_0}(m) \notin S$ (נזכיר כי $S = \{E_k(m_0) \,|\, k \in K\}$), ולפיכך קיים m_1 שעבור מחצית מהמפתחות מקיים $E_k(m_1) \notin S$.
ולכן המרחק בין ההתפלגויות $E_K(m_0)$ ו $E_K(m_1)$ הוא לפחות חצי, כלומר $\Delta(E_K(m_0), E_K(m_1)) \geq \frac{1}{2}$ ולכן ההסתברות להצלחה במשחק ההבחנה היא לפחות $\frac{1}{2} + \frac{1}{2}\Delta = \frac{3}{4}$. **הוכחה:** (לפי ברק, מעט יותר פורמאלי) נניח כי המפתחות באורך $n-1$. יהי $x_1 = 0^n$. נגדיר $S = supp(E_{U_{n-1}}(x_1))$. נשים לב כי מכיוון ש-$|K| = 2^{n-1}$ אזי $|S| \leq 2^{n-1}$. נגדיר את 2^{n-1} המשתנים המקריים $T_k = 1_{E_k(x)=1}$. (נשים לב כי משתנים מקריים אלו מתפלגים על ערכי ההודעות x). מכיוון שהצפנה חייבת לקיים חח"ע (על מנת שנוכל לשחזר הודעות) ומכיוון ש- $|S| \leq 2^{n-1} = |M| \geq 2^{n-1}$ מתקיים לכל k $Pr[T_k = 1] \leq \frac{1}{2}$. נשים לב כי T_k הינו משתנה מציין ולכן תוחלתו שווה להסתברות עצמה. נגדיר $T = \Sigma_{k \in \{0,1\}^{n-1}} T_k$. מתקיים

$$\mathbb{E}[T] = \Sigma_{k \in \{0,1\}^{n-1}} E[T_k] \leq 2^{n-2}$$

ומכאן נובע כי

$$Pr[T \leq 2^{n-2}] > 0$$

כלומר, קיים x כך ש- $\Sigma_{k \in \{0,1\}^{n-1}} T_k(x) \leq 2^{n-2}$. ובמילים, קיים x_2 כך לכל היותר מחצית המפתחות k מקיימים $E_k(x_2) \in S$. מכאן, $Pr[E_{U_n}(x_2) \in S] \leq \frac{1}{2}$. מנגד, $Pr[E_{U_n}(x_1) \in S] = 1$ לפי הגדרה. מכאן נובע $\Delta(E_{U_n}(x_1), E_{U_n}(x_2)) \geq \frac{1}{2}$. לפי משפט קודם סיכויי ההצלחה במצב זה הינם $\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$. ■

אם כך הראינו כי הדרישה החלפוית לבטיחות סטטיסטית לא אפשרה לנו להפחית בדרישה על אורך המפתחות.
שאלה - כאשר משתמשים ב OTP, רשימה חד פעמית, האם השיפור הבא עוזר: אם קיבלנו במקרה סדרה k שכולה אפסים, נחזור ונבחר משהו אחר?

2.2 $Message\ Authentication$

הערה 2.16 רקע מומלץ לחלק זה - http://cseweb.ucsd.edu/users/mihir/cse207/w-mac.pdf

עבור רבים, פרטיות היא המטרה העיקרית הקשורה להצפנה. עם זאת, מזווית מעט שונה נראה מעט הגיוני יותר לטעון כי אימות מידע היא מטרה חשובה יותר. ובאמת, לרב פחות יפריע לאדם לחשוף את פרטיהם (למשל ב-$facebook$) אך אותו אדם מאד לא ירצה שמישהו יוכל להתחזות לו ולהעביר הודעות בשמו.

אם כן, אנו עוברים לדבר מעט על סביבה מעט שונה. עד כה המוטיבציה של בוב ואליס היתה להעביר בינהם הודעות כך שמאזין (איב) לא תוכל לפענח את תוכן ההודעה. כעת, נדון במקרה בו בוב ואליס אדישים ליכולת של איב לפענח את תוכן ההודעות המועברות אך מעוניינים שאיב לא תוכל בעצמה לשדר הודעה מוצפנת ע"פ אותו המנגנון.

תאור הסביבה: לאליס ולבוב יש מפתח סודי משותף $k \in \{0,1\}^n$.

נדון בשתי שאלות:

1. אליס שולחת הודעה לבוב והם רוצים, בעזרת המפתח המשותף, להוסיף להודעה תוספת כך שצד שלישי לא יוכל לייצר הודעות חוקיות אחרות.
2. נדון בהמשך.

עבור השאלה הראשונה - נגדיר:

$$E : K \times M \to \{0,1\}^s$$

כך ש:

- קל לחשב את E
- גם אם רואים הודעה $(m, E_k(m))$, אזי ההסתברות למצוא $m' \neq m$ ו r' כך ש $E_k(m') = r'$ תהיה קטנה.

נעשה שימוש בפונקציות עירבול $(Hashing)$

נזכיר כי אוסף פונקציות H כך שלכל $h \in H$ מתקיים $h : \{0,1\}^n \to \{0,1\}^s$ נקרא אוסף 2 - אוניברסלי אם:

1. לכל $m \in \{0,1\}^n$אם בוחרים $h \in H$ מקרי, אז $h(x)$ מפולג באופן אחיד על $\{0,1\}^s$.
2. לכל $m_1 \neq m_2 \in M$ המשתנים המקריים $h(m_1)$ ו $h(m_2)$ עבור h מקרי, הם בלתי תלויים.

הערה 2.17 תזכורת נוספת: H אוניברסלי אם לכל $m_1 \neq m_2$ אם

$$Pr_{h \in H}[h(m_1) = h(m_2)] \leq \frac{1}{2^s}$$

נגדיר את $E : H \times \{0,1\}^n \to \{0,1\}^s$ ע"י $E_h(m) = h(m)$.

ואז למעשה ההודעה שנשלח תהיה ההודעה עצמה והעירבול שלה, דהיינו $(m, h(m))$. כאשר h נבחר באופן מקרי מ H 2 אוניבסלית ואינו ידוע ל E (זהו המפתח הסודי שלנו).

בגלל ה 2 - אוניברסליות, אחרי שאיב ראה הודעה אחת ואת קידודה, לכל הודעה אחרת m_2 - על פי תכונה 2 - $h(m_2)$ מתפלג עדיין באופן אחיד, ולכן ההסתברות להצלחה של איב היא $\frac{1}{2^s}$.

2.3 תרגיל - סודיות סטטיסטית

ל-A, B מפתח משותף סודי $k \in \{0,1\}^n$. אינפורמציה כלשהי אודות k דלפה לאיב. לצורך העניין איב יודעת כי $k \in S$ כך ש-$S \subseteq \{0,1\}^n$. נניח כי $|S| = 2^{n/2}$. A, B יודעים את גודל $|S|$ אך אינם יודעים את זהות האיברים ב-S. כעת הם רוצים לבחור מפתח חדש באמצעות הודעות בגודל $m < n$ אשר יתכן ϵ- בטיחות סטטיסטית עבור הודעות באורך m. לו יבחרו פונקציה מקרית $R : \{0,1\}^n \to \{0,1\}^{n/8}$ אז $R(k)$ יהיה מפתח חדש ונשים לב כי $R(S)$ מתפלג מקרית. מבחינה מעשית, בניית פונ' מקרית אינה יעילה (יש להתאים לכל איבר במקור שגודלו 2^n איבר בתמונה). במקום זאת בוב ואליס יכולים לקחת פונ' $hash$ מקרית מתוך אוסף 2-אוניברסלי ולקבוע $r = h(k)$ בתור מפתח חדש. במקרה זה ניתן להוכיח (וזה יהיה התרגיל) שבהסתברות אקס' קטנה (ב-n), אליס ובוב יקבל הצפנה ϵ-בטוחה עבור ϵ אקספוננציאלית קטן.

2.4 דוגמאות לפונקציות $Hash$

נראה מספר דוגמאות לפונקציות $Hash$ מ־$\{0,1\}^n$ ל־$\{0,1\}^m$ עבור $m \leq n$.

1. ניקח מטריצה $A \in M_{m\times n}(Z_2)$ ווקטור $b \in \{0,1\}^n$. נביט בפונ' $h_{A,b}(x) = (Ax+b) mod\, 2$. נגדיר $H = \{h_{A,b}\}$. חסרונות - צריך $m \times n$ ביטים על מנת לשמור את איברי המטריצה.

2. בוחרים סדרת ביטים $b = b_1, ..., b_n, ..., b_{n+m-1}$ וסדרה ביטים נוספת $c = c_1, ..., c_m$. מגדירים $y_j = b_j \cdot x_1 + ... + b_{n+j-1} \cdot x_n$ לכל $j \in [m]$. ואז אנו נגדיר $h_{b,c}(x) = y + c$. כאן אנו זקוקים לביטים בסדר גודל $O(n+m)$ וזה כמובן חסכון משמעותי לעומת הדוגמא הקודמת.

3 $Computational\ models\ and\ computational\ security.$

הראינו כי גם תחת ההגדרות של בטיחות סטטיסטית איננו יכולים להצפין הודעות באורך קטן מאורך ההודעות. במציאות, אנשים משתמשים במפתחות באורכים קצרים מאורכי ההודעות המקוריים עבור אפליקציות שדורשות זהירות רבה (אפליקציות שמעבירות מספרי כרטיס אשראי למשל). האם נוכל להשתמש בהוכחות מפרק קודם על מנת לפרוץ את המערכות הללו?

למעשה, אותן הוכחות מאפשרות לנו לכתוב קוד (קצר מאד) שפורץ מערכות כללו. הבעיה היא מעשית: עבור ערכי n גדולים, זמן הריצה של התוכנית יהיה בסדר גודל של 2^n. אפילו עבור ערכי n לא גדולים במיוחד, מדובר בזמן ריצה גדול במיוחד. למשל, אם ניקח מפתח באורך $1Kbit$, הרצת 2^{1000} פעולות ימשכו לנצח. מסתבר שזה בדיוק הרעיון שעומד מאחורי הצפנות עם מפתחות קצרים מאורך ההודעות - פיתוח מערכות הצפנה שאינן ניתנות לשבירה "בזמן סביר".

נרצה כעת לנסות ולתת הגדרות מדויקות שיתארו מערכת הצפנה שאיננת ניתנת לשבירה ב"זמן סביר". כמובן שברצוננו לספק תוצאות אשר יהיה להם תוקף ללא תלות בסוג האמצעים העומדים לרשת איב. לשם כך, לא נשאל איזה מעבד, מערכת הפעלה עומדים לרשות איב. במקום זאת ניעזר במודלים חישובים (שקולים) כגון מעגלים בוליאנים ומכונות טיורינג.

אנו נרשה מספר דברים:

1. האלגוריתמים שנפעיל יכולים להיות הסתברותיים.

2. נרשה לאלגוריתמים שלנו לקבל מידע נוסף (בצורת $string$) (זה לא הכרחי אך מפשט כמה מההוכחות שניתן).

אנו נדון בפונקציות חישוב f המקבלות קלט באורך n ביטים ונאמוד את זמן החישוב של f. על מנת להראות כי f ניתנת לחישוב בזמן T נוכל להראות באופן שקול את אחת התכונות הבאות:

1. f ניתנת לחישוב ע"י מעגל בוליאני בעל T שערים לוגיים.

2. f ניתנת לחישוב ע"י מכונת טיורינג הסתברותית שתיאורה אינו עולה על T תווים.

3. f ניתנת לחישוב ע"י תוכנית (בשפת c למשל) בעלת T תווים לכל היותר אשר עוצרת אחרי T צעדים לכל קלט.

3.1 התפלגויות בלתי ניתנות להבחנה חישובית

הגדרה 3.1 פונקציה $\epsilon : N \to [0,\infty)$ תקרא זניחה פולינומיאלית אם לכל $k \geq 0$ מתקיים $n^k \cdot \epsilon(n) \to 0$ עבור $n \to 0$.

דוגמאות: $\epsilon(n) = 2^{-n}$, $\epsilon(n) = \frac{1}{n^{loglogn}}$, $\epsilon(n) = 2^{-\sqrt{n}}$.

אנטי דוגמא: $\frac{1}{n^{64}}$.

הגדרה 3.2 פונקציה $\epsilon : N \to [0,\infty)$ תיקרא חסומה פולינומיאלית אם $\epsilon(n) \geq \frac{1}{n^{O(1)}}$.

דוגמאות: $\epsilon(n) = \frac{1}{10}, \epsilon(n) = \frac{1}{n^2}$

הערה 3.3 מעתה לעיתים נדבר על פונקציות ולא נטרח לציין כי מדובר בפונ' פולינומיאליות. אם לא יצויין אחרת נבין כי מדובר בפונקציות פולינומיאליות. כמו כן לעיתים נכתוב רק ϵללא ה־n כאשר זה יהיה ברור מההקשר.לבסוף, הקונבנציה עליה נסכים היא כי פוילנומיאלי $\equiv$יעיל.

הגדרה 3.4 יהיו $X = \{X_n\}, Y = \{Y_n\}$ (עבור n טבעי), סדרת התפלגויות, קרי X_n, Y_n התפלגויות על $\{0,1\}^{l(n)}$ כאשר $l(n) = poly(n)$. נאמר ש**לא ניתן להבחין סטטיסטית** בין X, Y אם $\Delta(X_n, Y_n)$ זניחה (לכל פונקציה חסומה פולינומיאלית ϵ קיים N_0 כך שלכל $n > N_0$ מתקיים $\Delta(X_n, Y_n) \leq \epsilon(n)$). במצב זה, נסמן $X \equiv Y$.

ניתן כעת הגדרה שקולה להגדרה האחרונה.

הגדרה 3.5 היו $X = \{X_n\}, Y = \{Y_n\}$ (עבור n טבעי), סדרת התפלגויות, קרי X_n, Y_n התפלגויות על $\{0,1\}^{l(n)}$ כאשר $l(n) = poly(n)$. נאמר שלא **ניתן להבחין סטטיסטית** בין X, Y אם לכל פונ' $A = \{A_n : \{0,1\}^{l(n)} \to \{0,1\}\}$ מתקיים:

$$|Pr_{z \sim X_n}[A_n(z) = 1] - Pr_{z \sim Y_n}[A_n(z) = 1]$$

זניח (זניח - כמו מקודם).

הגדרה 3.6 נאמר ש**לא ניתן להבחין חישובית** בין $X = \{X_n\}, Y = \{Y_n\}$ כאשר X_n, Y_n התפלגויות על $\{0,1\}^{l(n)}$ ו-$l(n) = poly(n)$ אם לכל אלגוריתם הסתברותי פולינומיאלי (ב-n) A מתקיים:

$$|Pr_{z \sim X_n}[A_n(z) = 1] - Pr_{y \sim Z_n}[A_n(z) = 1]|$$

זניח (כמו מקודם). במצב זה נסמן $X_n \approx Y_n$.

הערה 3.7 אם נרצה להיות קפדניים ולהתחשב במקרה בו $l(n)$ קטן בצורה משמעותית מ-n, נוכל לומר שאנו מעבירים את 1^n (נזכור כי מותר לנו לקודד מידע נוסף עבור האלגוריתם) כקלט נוסף על מנת להבטיח שהאלגוריתם רץ בזמן פולינומיאלי. המקרה המעניין הוא כאשר $l(n) > n$ ולכן לא נקפיד בעניין זה.

משפט 3.8 נציין מספר תכונות של התפלגויות בלתי ניתנות להבחנה חישובית.
1. סימטריות: $Y_n \approx X_n \Leftarrow X_n \approx Y_n$
2. $\Delta(X_n, Y_n) \leq \epsilon(n)$ עבור $\epsilon(n)$ זניחה $\Leftarrow X_n \approx Y_n$.
3. טרנזיטיביות: $X_n \approx Z_n \Leftarrow X_n \approx Y_n, Y_n \approx Z_n$
4. $X_n \approx Y_n$ וגם f פונקציה הניתנת לחישוב בזמן פולינומיאלי אזי $f(X_n) \approx f(Y_n)$.
5. $X_n \approx Y_n$ אזי לכל $m < n$ הרישות בגודל m של של X_n, Y_n אינן ניתנות להבחנה.

הוכחה: 3: נשתמש באי שיוויון המשולש לפיו $|a + b| \leq |a| + |b|$. מתקיים:

$$|Pr[A(X_n) = 1] - Pr[A(Y_n) = 1]| \leq |Pr[A(X_n) = 1] - Pr[A(Y_n) = 1]| + |Pr[A(Y_n) = 1] - Pr[A(Z_n) = 1]| \leq 2\epsilon(n)$$

הוכחת שאר הסעיפים: תרגיל. ■

3.1.1 טרנזיטיביות פולינומיאלית

אנו יכולים להרחיב את תכונה 3 (טרנזיטיביות).

משפט 3.9 נניח עבור $X^1, ..., X^m$ התפלגויות כך ש-$m = poly(n)$ ומתקיים $X^i \approx X^{i+1}$ לכל i. אזי, $X^i \approx X^m$.

הוכחה: כמו מקודם ההוכחה נובעת מוואריציה מוכללת של אי שוויון המשולש לפיה $|\sum_{i=1}^{m} a_i| \leq \Sigma_{i=1}^{m} |a_i|$. אנו מסיקים לפי אי שיוון זה כי מתקיים

$$|Pr[A(X_1) = 1] - Pr[A(X_m) = 1]| \leq m \cdot \epsilon$$

מכיוון ש-ϵ זניחה ו-$m = poly(n)$ אזי גם $m \cdot \epsilon$ זניחה. ■

3.2 הצפנה בטוחה חישובית

הגדרה 3.10 הצפנה **בטוחה חישובית:** מערכת הצפנה $(E,D)=(E_n,D_n)$ כאשר

$$E_n:\{0,1\}^n\times\{0,1\}^{l\{n\)}\to\{0,1\}^{k(n)},\ D_n:\{0,1\}^{k(n)}\to\{0,1\}^{l(n)}$$

כך ש־

$$D_n\circ E_n(m)=m$$

ו־l,k פונקציות פולינומיאליות, תיקרא בטוחה חישובית אם לכל זוג הודעות $m_0,m_1\in\{0,1\}^{l(n)}$ (באופן פורמאלי יותר נכון לדבר על זוג סדרות של הודעות), לא ניתן להבחין חישובית בין $X_n=E(U_n,m_0),Y_n=E(U_n,m_1)$.

הערה 3.11 הגדרה בעזרת משחק כנגד איב שקולה. אמנם בהגדרה הקודמת אנו דורשים קיום לכל זוג סדרות m_0,m_1 ובמשחק איב בוחרת בעצמה. קיימת אפשרות שקיימות סדרות הודעות שלא מקיימות את התנאים איך איב תתקשה לאתר אותן. אך איב יכולה לקבל את m_0,m_1 כקלט נוסף (אמרנו שניתן לקודד $advice$ ־ קלט נוסף) למרות שאיב עצמה פולינומיאלית.

תרגיל: הוכיחו כי ההגדרה הבאה עבור הצפנה בטוחה חישובית הינה הגדרה שקולה:

מערכת הצפנה $(E,D)=(E_n,D_n)$ כאשר

$$E_n:\{0,1\}^n\times\{0,1\}^{l\{n\)}\to\{0,1\}^{k(n)},\ D_n:\{0,1\}^{k(n)}\to\{0,1\}^{l(n)}$$

כך ש־

$$D_n\circ E_n(m)=m$$

ו־l,k פונקציות פולינומיאליות, תיקרא בטוחה חישובית אם לא ניתן להבחין חישובית בין $X_n=E_n(U_n,m),Y_n=E_n(U_n,U_{l(n)})$ לכל $m^n\in\{0,1\}^{l(n)}$.

פתרון: לשם הנוחות נוכיח על הודעות ולא על סדרות של הודעות.

מתקיים $Pr[A(Y_n)=1]=\mathbb{E}_{m\sim U(l(n))}[Pr[A(U_n,m)=1]$. נסמן הסתברות זו ב־$q$. כמו כן, נסמן ב־$p$ את ההסתברות $Pr[A(X_n)=1]$.

הגדרה 1 $\Leftarrow$הגדרה 2: נניח בשלילה כי $|p-q|>\epsilon$. אזי מתכונת התוחלת קיים m עבורו נקבל כי ההגדרה המקורית אינה מתקיימת.

הגדרה 2 $\Leftarrow$הגדרה 1: נניח כי לכל m_0 מתקיים $|p_{m_0}-q|\leq\epsilon$. אזי לכל m_0,m_1 מתקיים $|P_{m_o}-p_{m_1}|\leq 2\cdot\epsilon$.

3.3 יוצרים פסאודו־אקראיים

הגדרה 3.12 (ברק) אנו נאמר שההתפלגות $\{X_n\}$ היא פסאודו אקראית אם לא ניתן להבחין בינה לבין ההתפלגות האחידה U_n.

הגדרה 3.13 תהי פונקציה $G:\{0,1\}^n\to\{0,1\}^{l(n)}$. G תיקרא יוצרת פסאדו־אקראית אם מתקיים לגביה:

1. G ניתנת לחישוב פולינומיאלי.
2. $l(n)>n$.
3. לא ניתן להבחין חישובית בין $G(U_n)$ לבין $U(l(n))$.

ברור על פי תכונות שהראינו קודם כי טריוויאלי לייצר G עבור $l(n) = n$ (נניח העתקת הזהות). כמו כן, אם קיימת G עבור $l(n)$ כלשהו אזי ניתן לייצר (ע"י קיטום לפי תכונה 5) G עבור $l'(n)$ לכל $l'(n) < l(n)$. מסתבר שקיימות תוצאות מרשימות בהרבה.

משפט 3.14 (**אקסיומת PRG**) קיים יוצר פסאודו אקראי עבור $l(n) = n + 1$.

הערה 3.15 אולי נוח יותר לחשוב על אקסיומת PRG בתור השערה. משערים כי קיים יותר פסאודו אקראי כנ"ל.

משפט 3.16 אם אקסיומת PRG נכונה אזי לכל קבוע c קיימת מערכת הצפנה בטוחה חישובית עבור הודעות באורך $l(n) = n^c$.

נכונות המשפט הנ"ל תנבע מנכונות המשפטים הבאים:

משפט 3.17 אם יש יוצר פסאודו אקראי G עם מתיחה $l(n)$ אזי קיימת מערכת הצפנה בטוחה חישובית (E, D) שבעזרתה ניתן להצפין באמצעות מפתחות באורך n הודעות באורך $l(n)$.

הוכחה: יהי G יוצר פסאדו אקראי הממפה מחרוזות באורך n ביטים למחרוזות באורך $l(n)$ ביטים. נגדיר את מערכת ההצפנה כדילקמן:

$$E_k(x) = x \oplus G(k) =, D_k(y) = y \oplus G(k)$$

קל לוודא תקינות: $D_k(x \oplus G(k)) = x \oplus G(k) \oplus G(k) = x$.
כעת מספיק להראות שההתפלגות $E_{U_n}(x)$ היא פסאודו־אקראית לכל הודעה x. אם נראת זאת אזי לכל זוג הודעות x^0, x^1 מתקיים

$$E_{U_n}(x_0) \approx U_{l(n)}, E_{U_n}(x_1) \approx U_{l(n)}$$

ומטרנזיטיביות נקבל $E_{U_n}(x_0) \approx E_{U_n}(x_1)$.
אם כן, נניח בשלילה כי התפלגות זו איננה פסאודו אקראית. אז קיים אלגוריתם פולינומיאלי A כך שמתקיים

$$|Pr[A(G(U_n) \oplus x) = 1] - Pr[A(U_{l(n)}) = 1] \geq \epsilon \qquad *$$

נגדיר כעת אלגוריתם $B : \{0,1\}^{l(n)} \to \{0,1\}$ באופן הבא:

$$B(y) = A(y \oplus x)$$

נשים לב כי לפי הגדרה זו $A(z) = B(z \oplus x)$. זמן הריצה של B זהה לזה של A. לפי * נסיק:

$$|Pr[B(G(U_n)) = 1] - Pr[A(U_{l(n)} \oplus x) = 1] \geq \epsilon$$

מכיוון ש־$U_{l(n)} \oplus x \equiv U_{l(n)}$ (נובע מכך ש־xor פונקציה חח"ע ועל) אנו מקבלים אלגוריתם (B) אשר מבחין בין $G(U_n)$ ל־$U_{l(n)}$ בסתירה להנחה. ■

משפט 3.18 אם קיים יוצר פסאודו אקראי עבור $l(n) = n + 1$ אז קיים יוצר פסאודו אקראי עבור $l(n) = n^c$ לכל קבוע c.

הוכחה: נניח כי קיים יוצר פסאודו אקראי pmG הממפה n ביטים ל־$n+1$ ביטים. ניצור באמצעותו יוצר פסאודו אקראי G^m הממפה n ביטים ל־$l(n)$ ביטים. זמן הריצה של האלגוריתם יהיה $l(n)$ פעמים $\times$זמן הריצה של של pmG. נסמן $x_{[i,...,j]}$ להיות המחרוזת $x_i...x_j$. כמו כן נסמן
האלגוריתם:

קלט: $x \in \{0,1\}^n$
$j \leftarrow 0$
$x^{(0)} \leftarrow x$
כל עוד $j < l(n)$
$j \leftarrow j+1$
$x^{(j)} = pmG(x^{(j-1)}_{[1,...,n]})$
$output\, x^{(j)}_{[n-1]}$

נגדיר את ההתפלגויות $Y^{(0)}, Y^{(1)}, ..., Y^{(m)}$ (m הרלוונטי במקרה שלנו הוא $l(n)$) על $\{0,1\}^m$ בצורה הבאה: Y_i תתאר את ריצת האלגוריתם החל מהאיטרציה ה־i על ההתפלגות U_{n+i}. בצורה פורמאלית יותר, Y_i מתקבלת משרשור i ביטים אקראיים לריצת G^{m-i} (הרצת האלגוריתם למשך $m-i$ איטרציות) על קלט אקראי מתוך U_n, משמע $Y_i = U_i G^{m-i}(U_n)$. נשים לב כי $Y^{(0)} = G^m(U_n)$ ו־$Y^{(m)} = U_m$. נותר להראות כי $Y^{(m)} \approx Y^{(0)}$. לפי תכונת הטרנזיטיביות הפולינומיאלית, מספיק שנראה כי לכל $i \in [m]$ מתקיים $Y_i \approx Y_{i+1}$. אם כן, נשים לב כי $Y^{(i)} = U_i G^{m-i}(U_n), Y^{(i+1)} = U_{i+1} G^{m-i-1}(U_n)$. מכיוון שלשני המשתנים המקריים האלה ריישא של i ביטים אקראיים, מספיק שנראה $X = G^{m-i}(U_n) \approx Y = U_1 G^{m-i-1}(U_n)$. נגדיר $f : \{0,1\}^{n+1} \to \{0,1\}^{m-i}$ בצורה הבאה:

$$f(y) = y_{[n+1]} G^{m-i-1}(y_{[1,...,n]})$$

(נשים לב כי f פולינומיאלית)
נשים לב כי: $X = f(pmG(U_n)), Y = f(U_{n-1})$. אך מכיוון ש־$pmG(U_n) \approx U_{n+1}$ אנו מקבלים לפי משפט שהוכחנו כי

$$X = f(pmG(U_n)) \approx f(U_{n-1}) = Y$$

כנדרש. ■

3.3.1 דוגמאות ליוצרים פ"א

הערה 3.19 כאמור אין הוכחה לקיומו של יוצר פ"א. כאשר אנו מספקים דוגמא ליוצר פ"א, אנו מתכוונים שאנו משערים שהפונקציה המדוברת מהווה יוצר פ"א.

ניתן כעת דוגמא ליוצר פ"א.
נגדיר פונקציה $G : \{0,1\}^n \to \{0,1\}^m$ באופן הבא: בוחרים שני מספרים ראשוניים (ע"י שימוש באלגוריתם ההסתברותי של רבין למשל) p, q בני $n/2$ ביטים. דורשים גם $p \equiv q \equiv 3 (mod\, 4)$ קובעים $N = p \cdot q$. בוחרים מספר $x \in \{0,1\}^n$. כעת מגדירים עבור $k \in [m-1]$:

$$x_1 = x, ..., x_{k+1} = x_k^2 (mod\, N)$$

וכמו כן מגדירים עבור $k \in [m]$

$$b_k = lsb(x_k)$$

לבסוף מגדירים $G(x) = b_1...b_m$.

משפט 3.20 אם A אלגוריתם ϵ־מבחין בין $G(U_n)$ לבין U_m ונסמן את זמן הריצה של A ב־T אז יש אלג' פולינומיאלי ב־$(n, m, \frac{1}{\epsilon}, T)$ שמפרק את N לגורמים.

הוכחה: לא תינתן. "זה ארוך ומסורבל". ■

הערה 3.21 N הוא בסדר גודל של 2^n והאלגוריתם היעיל ביותר לפירוק מספר בסדר גודל כזה לגורמים רץ בזמן $2^{c \cdot n^{1/3} \cdot log(n)^{2/3}}$.

על פניו, נראה ש־G יכול לסייע לנו. אולם, על מנת להשיג חסמים טובים באמצעות הרדוקציה שהראינו קודם עלינו להכפיל סדר גודל של 2000 ביטים. זה לא כל כך יעיל. נרצה לתאר מנגנונים נוספים. על מנת לעשות כן, נגדיר תחילה את המושג פונקציה חד־כיוונית. באופן לא פורמאלי פונקציה חד־כיוונית היא פונקציה הקלה לחישוב וקשה "להפוך". תחילה ניתן הגדרה (ע"פ $Trevisan$) עבור פונקציה סופית.

3.3.2 פונקציות חד־כיווניות

הגדרה 3.22 פונקצייה חד־כיוונית עבור $S \in \mathbb{N}$ ועבור $0 < \epsilon < 1$, נאמר כי $f : D \to R$ היא פונקציה (s, ϵ)־חד־כיוונית ביחס ל־$Domain$ שנסמנו X, אם לכל מעגל C בגודל S לכל היותר (שמנסה להפוך את f), מתקיים

$$Pr_{x \sim \{0,1\}^n}[C(f(x) = x' \; s.t \; f(x) = f(x')] \leq \epsilon$$

כאשר ההסתברות היא על פני בחירה יונפורמית של x מההתפלגות X.

הערה 3.23 כמו במקרים רבים, לעיתים נסתפק בנוטציה אסימפטוטית ולא נתייחס לפרמטרים S, ϵ באופן ספציפי.

הגדרה 3.24 פונקציה חד־כיוונית - הגדרה אסימפטוטית פונקציה $f : \{0,1\}^* \to \{0,1\}^*$ היא חד כיוונית אם:
1. f ניתנת לחישוב בזמן פולינומיאלי.
2. לכל פולינום $t()$ קיימת פונקציה זניחה v כך שלכל n גדול מספיק מתקיים $f_n(x) = (n, f(x))$ היא $(t(n), v(n))$ חד כיוונית.

הגדרה 3.25 פרמוטציה חד כיוונית אם $f : \{0,1\}^n \to \{0,1\}^n$ פונקציה חח"ע (t, ϵ) חד־כיוונית, אז נקרא ל־f פרמוטציה (t, ϵ) חד־כיוונית.
דוגמאות לפונקציות חד כיווניות:

הערה 3.26 כפי שהערנו לגבי יוצרים פ"א - אנו מניחים שקיימות פונקציות חד־כיווניות אך אין הוכחה לכך. מאחורי כל דוגמא לפונקציה חד כיוונית מסתתרת ההשערה שקיימות כאלה בכלל.

דוגמא א': יהי p ראשוני בן n ביטים. אזי נסתכל על החבורה $Z_p^* = \{1, ..., p-1\}$ (כאשר הפעולה המוגדרת עליה היא כפל מודלו p). כפי שלמדנו בעבר, זוהי חבורה ציקלית. לכן, קיים יוצר g כך ש־

$$1, g^1, g^2, ..., g^{p-2}$$

כולם שונים (כלומר פורשים את איברי החבורה). כמו כן, לכל $x \in Z_p^*$ מתקיים $x^{p-1} = 1$. נגדיר $f : \{0, ..., p-2\} \to \{1, ..., p-1\}$ כך:

$$f_{g,p}(x) = g^x (mod\, p)$$

f הינה חח"ע ולכן מהווה פרמוטציה.
עבור רב הבחירות של p, g אנו משערים כי כי f הינה פונקציה חד־כיוונית. האלגוריתם הטוב ביותר (הידוע) להפיכת הפונקציה רץ בזמן $2^{O((logp)^{1/3})}$.

דוגמא ב' (Subset Sum): עבור קלט $x \in \{0,1\}^n$, עבור $n = k \cdot (k+1)$, $SS_k(x)$ מפרק את x ל־k מספרים, כל אחד באורך k ביטים ותת־קבוצה $I \subseteq \{1, ..., k\}$. הפלט הינו

$$SS_k(x_1,...,x_k,I) = x_1,...,x_k,\Sigma_{i\in I}x_i$$

אנו משערים כי SS_k היא פונקציה (t,ϵ)־חד כיוונית כאשר t,ϵ^{-1} סופר פולינומיאליים ב־k.

שאלה בהינתן פרמוטציה חד כיוונית $f:\{0,1\}^n\to\{0,1\}^n$ ידיעת $f(x)$ לא מספקת ידיעה באמצעותה ניתן לגלות את x. אך האם זה אומר כי לא מתקבל אף פרט על x?

תשובה לא בהכרח. אנו נראה כעת טענה לפיה נסיק שפונקציה חד כיוונית לא בהכרח מסתירה את כל המידע אודות x.

טענה 3.27 תהי $f:\{0,1\}^n\to\{0,1\}^n$ חד כיוונית. אזי גם $F:\{0,1\}^{2n}\to\{0,1\}^{2n}$ המוגדרת כך:

$$F(x,y) = (f(x),y)$$

גם כן חד כיוונית.

הוכחה: נניח בשלילה ש־F אינה חד כיוונית. אזי קיים אלגוריתם פולינומיאלי A' המוצא בהנתן $(f(x),y)$ כך x',y' ש־$y=y'$ וגם $f(x')=f(x)$ בהסתברות לא זניחה. אזי, ניתן להיעזר ב־A' על מנת להפוך את f (פשוט לקרוא ל־A' עם $f(x)$ ועוד n־יית ביטים כלשהי ולחלץ רק את x' הדרוש לנו) בסתירה לחד כיווניות של f. ■

כיצד הסקנו מטענה זו את תשובתנו לשאלה שהוצגה: הראינו פונקציה F חד־כיוונית, כך שלכל $z\in\{0,1\}^{2n}$ $F(z)$ מספקת מידע מלא על n הביטים הראשונים של z.

3.4 קיום פרמוטציה חד־כיוונית גורר קיום יוצר פ"א מרחיב

הגדרה 3.28 (ביט־קשה) תהי $f:\{0,1\}^n\to\{0,1\}^n$ פרמוטציה חד־כיוונית. פונקציה בוליאנית $B:\{0,1\}^n\to\{0,1\}$ הינה (t,ϵ) ביט־קשה עבור f אם לכל אלגוריתם A בעל סיבוכיות לכל היותר t מתקיים

$$Pr_{x\sim U_n}[A(f(x))=B(x)]\le\frac{1}{2}+\epsilon$$

הערה 3.29 כמו תמיד, קיימת ההגדרה האסימפטוטית המתעלמת מהפרמטרים (s,ϵ) ודורשת ש־B פולינומיאלית וכמו כן דורשים את קיום האי שיוון עבור n מספיק גדול (עבור $\epsilon(n)$ זניחה במקום ϵ קבוע).

טענה 3.30 נניח כי B היא ביט קשה עבור פרמוטציה $f:\{0,1\}^n\to\{0,1\}^n$ אזי f פרמוטציה חד־כיוונית.

הוכחה: נניח בשלילה כי קיים אלגוריתם פולינומיאלי C ההופך את f, כלומר לכל פונקציה זניחה $\epsilon(n)$ מתקיים:

$$Pr_{x\sim U_n}[C(f(x))=x]>\epsilon(n)$$

(לשם הנוחיות נעבוד מכאן עם ϵ קבוע)
נציע אלגוריתם A שיפעל כך: בהנתן $f(x)$ יפעיל את C ויקבל פלט y. בודקים האם $f(y)=f(x)$ (מאורע שמתרחש בהסתברות $p>\epsilon$). אם כן, $y=x$ (כי f פרמוטציה) ואז נחשב $B(x)$ וזה יהיה הפלט. אחרת, נטיל מטבע וזה יהיה הפלט. מתקיים:

$$Pr_{x\sim U_n}[A(f(x))=B(x)] = p+\frac{1}{2}(1-p) = \frac{1}{2}+\frac{p}{2} > \frac{1}{2}+\frac{\epsilon}{2}$$

בסתירה לכך ש־B ביט קשה. ■

הצצה להמשך אנו נראה כעת כי בהנתן פרמוטציה חד־כיוונית אנו יכולים לבנות B ביט־קשה עבור f. לאחר מכן, נראה כי בעזרת f פונקציה (ובפרט פרמוטציה) חד־כיוונית ו־B ביט־קשה עבור f ניתן לייצר יוצר פ"א הממפה l ביטים ל־$l+1$ ביטים.

אנו נראה כעת כי XOR מקרי הוא ביט קשה עבור כל פרמוטציה חד כיוונית.

הערה 3.31 כשנשתמש בסימון המפכלה הפנימית בחלק זה אנו נתכוון למכפלה הפנימית מודולו 2.

משפט 3.32 (Goldreich and Levin) נניח כי A אלגוריתם בסיבוכיות t כך ש־

$$Pr_{x,r}[A(f(x),r) =< x,r >] > \frac{1}{2} + \epsilon$$

אזי קיים אלגוריתם A' בסיבוכיות לכל היותר $O(t\epsilon^{-2}n^{O(1)})$ כך ש־

$$Pr_x[A'(f(x)) = x] > \Omega(\epsilon)$$

מסקנה 3.33 אם קיים אלגוריתם פולינומיאלי אשר סותר את היות XOR מקרי ביט קשה עבור פרמוטציה f, אזי f אינה חד־כיוונית.

נכונות המסקנה נובעת ישירות מן המשפט. על מנת להוכיח את משפט $G\&L$ נראה תחילה גירסה חלשה שלו.

משפט 3.34 ($\mathbf{Goldreich\ and\ Levin\ -\ Weak\ Version}$) נניח כי A אלגוריתם בסיבוכיות t כך ש־

$$Pr_{x,r}[A(f(x),r) =< x,r >] > \frac{15}{16} \qquad (3)$$

אזי קיים אלגוריתם A' בסיבוכיות לכל היותר $O(tn \cdot logn + n^2 \cdot logn)$ כך ש־

$$Pr_x[A'(f(x)) = x] > \frac{1}{3}$$

לפני שנוכיח את המשפט האחרון, יועיל לנו אם נחשוב שניה על גירסה עוד יותר חלשה. נניח שההסתברות ב־(3) היא 1 (ולא רק גדולה מ־$\frac{15}{16}$). במצב זה, הפיכת f הינה קלה מאד. נסמן ב־$e_i \in \{0,1\}^n$ את הווקטור בו בקורדינטה ה־i יש 1 ושאר הקורדינטות מאופסות. אזי $< x, e_i >= x_i$. כעת, בהינתן $y = f(x)$ ואלגוריתם A שעומד בתנאי (3) המשופרים (מקיים בהסבתרות 1), A' יריץ את A n פעמים. לכל $i \in [n]$ נריץ $A(y, e_i)$ ונקבל את x_i ולבסוף את x. כמובן ש־A' עומד בתנאי הסיבוכיות הנדרשים במשפט. על מנת להוכיח את הגירסה החלשה של $G\&L$ אנו נפעל בצורה דומה אך נאלץ להתמודד עם העובדה כי ההבטחה (המקורית) ב־(3) היא בהסתברות. אנו נקבל את $G\&L$ החלש מבניית האלגוריתם שנציג בלמה הבאה. עוד לפני שנציג את הלמה נזכיר למה המציגה חסם מוכר לריכוז מידה - חסם $Chernoff$ שיעזור לנו בהמשך.

למה 3.35 ($\mathbf{Chernoff\ Bound}$) יהיו $X_1, ..., X_n$ מ"מ ב"ת המתפלגים על התחום $[0,1]$. אזי, לכל $\epsilon > 0$מתקיים

$$Pr[\Sigma_{i=1}^n X_i > \mathbb{E}[\Sigma_{i=1}^n X_i] + \epsilon \cdot n] \leq e^{-2 \cdot \epsilon^2 \cdot n}$$

למה 3.36 (**G&L Algorithm – Weak Version**) קיים אלגוריתם GLW שבהינתן גישה לפונקציה $H : \{0,1\}^n \to \{0,1\}$ שמקיימת עבור $x \in \{0,1\}^n$ (מסויים)

$$Pr_{r \in U_n}[H(r) =< x, r >] > \frac{7}{8}$$

רץ בזמן $O(n^2 \cdot logn)$, מבצע $O(logn)$ שימושים ב־H ובהסתברות $1 - o(1)$ פולט את x.

הוכחה: תחילה הרעיון - אנו רוצים לחשב את $< x, e_i >$ לכל $i \in [n]$. עם זאת, איננו יכולים פשוט לחשב את $H(e_i)$ בגלל שיתכן ש־H טועה על קלטים כאלה (H "טובה" עבור קלט מקרי). אם נרצה לחשב $< x, r >$ עבור r מקרי נהיה במצב טוב יותר מפני ש־$H(r)$ תחשב עבורנו את הביטוי הנ"ל בהסתברות גבוהה. אם כן, נשתמש בזהות

$$< x, y >=< x, r + y > - < x, r >$$

ועל מנת לחשב $< x, y >$ (y במקרה שלנו זה e_i כלשהו), אנו יכולים לבחור r מקרי ולחשב $H(r+y) - H(r)$. אם r מתפלג באופן אחיד, גם $H(r+y), H(r)$ מפולגים באופן אחיד ומתקיים:

$$Pr_{r \sim U_n}[H(r+y) - H(r) =< x, y >] \geq Pr_{r \sim U_n}[H(r+y) =< x, r+y > \wedge H(r) =< x, r >]$$

$$= 1 - Pr_{r \sim U_n}[H(r+y) \neq< x, r+y > \vee H(r) \neq< x, r >] \geq$$

$$1 - Pr_{r \sim U_n}[H(r+y) \neq< x, r+y >] - Pr_{r \sim U_n}[H(r) \neq< x, r >] \geq 1 - 2 \cdot \frac{1}{8} = \frac{3}{4}$$

כעת, נניח כי אנו בוחרים $r_1, ..., r_k$ מקריים ומחשבים $Y_j = H(r_j + y) - H(r_j)$ לכל $j \in [k]$ וניקח את הערך (ביט) השכיח יותר. לפי הניתוח הקודם כל Y_j שווה ל־$< x, y >$ בהסתברות לפחות $3/4$. בנוסף, המאוראות $Y_j =< x, y >$ הם בת"ל. אנו יכולים להשתמש בחסם $Chernoff$ על מנת להסיק כי ההסתברות שהערך הנבחר יהיה שגוי הינה לכל היותר $e^{-k/8}$ (אם הערך שגוי זה אומר שעבור לפחות $k/2$ מה־Y_j קיבלנו שגיאה, למרות שבתוחלת היינו אמורים לקבל $k/4$. כלומר, קיבלנו סטיית תקן של $k/4$ מהתוחלת ולכן נשתמש בחסם עם $\epsilon = 1/4$ ונקבל את המבוקש).

האלגוריתם GLW: לכל $i \in [n]$ מחשב לכל $j \in [16logn]$ את $H(r_j + e_i) - H(e_i)$ (עבור r_j מקרי) וקובע את x_i לערך השכיח. לבסוף האלגוריתם מחזיר את x.

כעת נציב בניתוח הקודם $k = 16logn$ ונקבל כי לכל i ההסתברות לשגות בחישוב x_i הינה לכל היותר $e^{-2logn} = 1/n^2$. נשתמש בחסם האיחוד ונקבל כי ההסתברות לשגות בשיחזור x כולו היא לכל היותר $n \cdot \frac{1}{n^2} = \frac{1}{n} = o(1)$ כנדרש. כמו כן, האלגוריתם רץ בזמן $O(n^2 logn)$ ומבצע בסך הכל $32n \cdot logn$ שימושים ב־H. ■

הערה 3.37 מיכאל נכנס לנושאים בתורת הקודים בהוכחה האחרונה (הציג את ה־xor המקרי כקוד לתיקון שגיאות). העדפתי להסתמך על המקורות עליהם נשען מיכאל בהרצאה (הרצאותיו של $Trevisan$).

כעת נרצה להסיק את נכונות משפט $G\&L$ החלש מהלמה שהוכחנו. לשם כך נשתמש בחסם נוסף לריכוז מידה - אי־שיוויון מרקוב (ליתר דיוק, אנו נשתמש בווריאציה שלו)

למה 3.38 יהי X מ"מ (בדיד) מתפלג בתחום $[0,1]$. אזי לכל $0 \leq t \leq \mathbb{E}[X]$ מתקיים

$$Pr[X \geq t] \geq \frac{\mathbb{E}[X] - t}{1 - t}$$

נפנה כעת להוכחת משפט $G\&L$ החלש.
הוכחה: ההנחה במשפט יכולה להירשם כך:

$$\mathbb{E}_x[Pr_r[A(f(x),r) =< x,r >]] > \frac{15}{16}$$

(הסיבה לכך היא שהמ"מ הרלוונטי הוא משתנה מציין)
לפי אי שיוויון מרקוב (כפי שצוטט לעיל) נסיק:

$$Pr_x[Pr_r[A(f(x),r) =< x,r >] > \frac{7}{8}] > \frac{\frac{15}{16} - \frac{7}{8}}{\frac{1}{8}} = \frac{1}{2}$$

אנו נגיד ש־x "טוב" אם הוא מקיים $Pr_r[A(f(x),r) =< x,r >] > \frac{7}{8}$. האלגוריתם A', בהנתן קלט $y = f(x)$ יריץ את האלגוריתם GLW שתואר לעיל. אם x "טוב", לפי הלמה שהראינו, A' ישחזר את x בהסתברות לכל הפחות $1 - o(1)$. אנו יודעים לפי האי שיוויון הקודם כי לפחות מחצית ה־xים הם "טובים". לכן, בסך הכל, A' יצליח לשחזר את x עבור $\frac{1}{2} - o(1) > \frac{1}{3}$. זמן הריצה של A' הוא $O(n^2 \cdot logn)$. ■

נרצה להוכיח כעת את משפט $G\&L$ (החזק). מסתבר שאם נניח בלמת $G\&L\ Alogritm$ נכונות התנאי כשבצד ימין של אי השיוויון מופיע $\frac{3}{4} + \epsilon$ נוכל לקבל וריאציה של משפט $G\&L$ שגם כן דורשת קיום תנאי בהסתברות $\frac{3}{4} + \epsilon$ (במקום $\frac{1}{2} + \epsilon$). נראה אם כן כי אנו צריכים לדרוש בלמת $G\&L\ Alogritm$ נכונות התנאי בהסתברות $\frac{1}{2} + \epsilon$ נוכל להשיג את הנדרש. לרע מזלנו, זה לא יעבוד. מסתבר כי לכל $x, x' \in \{0,1\}^n$ ניתן לבנות H כך שמתקיים

$$Pr_{r \in U_n}[H(r) =< x,r >] = \frac{3}{4}$$

וגם

$$Pr_{r \in U_n}[H(r) =< x',r >] = \frac{3}{4}$$

ולכן אין אלגוריתם שיכול למצוא את x בהנתן H הנ"ל וזאת מפני ש־x לא מזוהה באופן יחיד ע"י H.
אנו נוכיח במקום זאת את הלמה הבאה:

למה 3.39 $(\mathbf{Goldreich - Levin\ Algorithm})$
נניח שיש לנו גישה לפונקציה $H : \{0,1\}^n \to \{0,1\}$ שעבור $x \in \{0,1\}^n$ כלשהו, מתקיים עבורה

$$Pr_{r \in U_n}[H(r) =< x,r >] \geq \frac{1}{2} + \epsilon$$

עבור $\epsilon > 0$ אז קיים אלגוריתם GL שרץ בזמן $O(n^2 \cdot \epsilon^{-4} \cdot logn)$ וקורא $O(n\epsilon^{-4}logn)$ קריאות לפונקציה H ומוציא רשימה $L \subseteq \{0,1\}^n$ כך ש־$|L| = O(\epsilon^{-2})$ ובהסבתרות לפחות $\frac{1}{2}$, $x \in L$.

הוכחה: אנו מקבלים $H()$ כך ש־$H(r) =< x,r >$ עבור לפחות $\frac{1}{2} + \epsilon$ מה־rים האפשריים. באמצעות H נרצה לבנות H' אשר 'מסכימה' עם $7/8$ מה־rים. ואז נוכל להשתמש באלגוריתם GLW ובתוצאות שהראינו לגביו על מנת למצוא את x. אנו מבצעים את הרדוקציה הזו ע"י 'ניחוש' הערך $< x,r >$בכמה נקודות שונות. לשם כך, אנו נבחר k נקודות אקראיות $r_1, ..., r_k \in \{0,1\}^n$ עבור $k = O(1/\epsilon^2)$. נניח לרגע שאנו יודעים את הערכים $< x,r_1 >, ..., < x,r_k >$. נוכל לקבוע את $H'(r)$ להיות הערך התדיר שמתבקל עבור $j = 1,2,...,k$

$$H(r + r_j) - < x,r_j >$$

לכל j, הביטוי הנ"ל שווה $< x, r >$ (כאמור, תחת ההנחה שאנו יודעים את ערכי $< x, r_j >$ לכל j) בהסתברות $\frac{1}{2} + \epsilon$ לפחות. ע"י בחירת k כנ"ל נוכל להבטיח

$$Pr_{r,r_1,...,r_k}[H'(r) =< x, r >] \geq \frac{31}{32}$$

לפי אותה וריאציה של א"ש מרקוב בה השתמשנו קודם מתקיים

$$Pr_{r_1,...,r_k}[Pr_r[H'(r) =< x, r >] \geq \frac{7}{8}] \geq \frac{3}{4}$$

מכאן מתקבל האלגוריתם הבא:

GL – first Attempt

בחר $r_{1,...,r_k} \in U_n$ עבור $k = O(1/\epsilon^2)$
לכל $b_1, ..., b_k \in \{0, 1\}$ (יש 2^k כאלה) בצע
הגדר $H'_{b_1...b_k}$ לפי הרב של $H(r + r_j) - b_j$
הפעל את אלגוריתם GLW עבור $H'_{b_1,...,b_k}$
הוסף את התוצאה לרשימה
החזר את הרשימה

הרעיון מאחורי האלגוריתם הזה הוא שאמנם איננו יודעים את הערכים $< x, r_j >$ לכל $j \in [k]$ אך ביכולתנו לנחש אותם ע"י מעבר על כל האפשרויות של בחירת ביטים b_j. לפי מסקנות קודמות, אם $H(r)$ מסכים עם $< x, r >$ עבור לפחות $\frac{1}{2} + \epsilon$ מערכי r, אזי קיימת הסתברות של לפחות $\frac{3}{4}$ שבאחת האיטרציות נריץ את GLW עם H' שמסכימה עם $< x, r >$ על לפחות $\frac{7}{8}$ ערכי r ואז על פי למת $GLW - Algorithm$ הרשימה הסופית תכיל את x בהסתברות של לפחות $3/4 - 1/n > 1/2$.

הבעיה באלגוריתם שהצגנו היא כמובן קביעת סדר הגודל של k. עבור k פולינומיאלי ב־$1/\epsilon$ האלגוריתם רץ (מספר האיטרציות) בזמן אקספוננציאלי ב־$1/\epsilon$. מסיבה זו, גם הרשימה שתיווצר הינה אקספוננציאלית ב־$1/\epsilon$.

על מנת להתגבר על הבעיה האחרונה נעבוד עם סדר גודל של $logO(1/\epsilon^2)$. מתקבל האלגוריתם הבא:

GL Algorithm

בחר $r_1, ..., r_t \in U_n$ כך ש־$t = logO(1/\epsilon^2)$
הגדר $r_S = \Sigma_{j \in S} r_j$ לכל תת קבוצה לא ריקה $S \subseteq \{1, ..., t\}$
לכל $b_1, ..., b_t \in \{0, 1\}$ (2^t כאלה) בצע
הגדר $b_S = \Sigma_{j \in S} b_j$ לכל תת קבוצה לא ריקה $S \subseteq \{1, ..., t\}$
הגדר $H'_{b_1,...,b_t}(r)$ לפי ערך תדיר על פני תתי הקבוצות $S \subseteq \{1, ..., t\}$ של הערכים $H(r + r_S) - b_S$
הרץ את האלגוריתם GLW עם $H'_{b_1,...,b_t}$
הוסף את התוצאה לרשימה
החזר את הרשימה

נבחן כעת מדוע האלגוריתם האחרון עובד. נשים לב כי בגלל ש־$r_1, ..., r_t$ אקראיים, גם הסכומים r_S, r_T אקראיים לכל תתי קבוצות $S \neq T$. כעת, יהי x כך ש־$< x, r >$ ו־$H(r)$ מסכימות על $\frac{1}{2} + \epsilon$ מה־rים. אז עבור בחירת $\{b_j\}$ כך ש־$b_j =< x, r_j >$ לכל j, מתקיים

$$b_S = \Sigma_{j \in S} b_j = \Sigma_{j \in S} < x, r_j >=< x, r_S >$$

לכל תת קבוצה לא ריקה S. במצב זה, לכל S ולכל r, בהסבתרות $\frac{1}{2} + \epsilon$ על פני בחירת r_j מתקיים

$$H(r + r_S) - b_S =< x, r >$$

וכמו כן האירועים הנ"ל בת"ל בזוגות. נסמן את $H'_{b_1,...,b_k}$ עבור $b_1,...,b_k$ האחרונים ב־H'.
נראה למה שתסייע לנו:
למה: יהיו $R_{1,...,}R_k$ קבוצה של מ"מ בת"ל בזוגות בתחום $0-1$. בנוסף נתון כי כל R_i שווה ל־1 בהסתברות לפחות $1/2+\epsilon$. אזי מתקיים

$$Pr[\Sigma_i R_i \geq k/2] \geq 1 - \frac{1}{4 \cdot \epsilon^2 \cdot k}$$

הוכחה: נסמן $R = R_1 + ... + R_k$. השונות של מ"מ בתחום $0-1$ היא לכל היותר $1/4$. כמו כן, בגלל אי התלות בזוגות, מתקיים:

$$Var[R] = \Sigma_i Var[R_i] \leq \frac{k}{4}$$

מכאן, מתקיים לפי אי שיוויון שביצ'ב

$$Pr[R \leq k/2] \leq Pr[|R - \mathbb{E}[R]| \geq \epsilon \cdot k] \leq \frac{Var[R]}{\epsilon^2 \cdot k^2} \leq \frac{1}{4\epsilon^2 k}$$

כנדרש.
נחזור להוכחת המשפט. בעזרת הלמה האחרונה נוכל לחסום הסתברותית את האפשרות שהערך התדיר שיתקבל עבור H' מחזיר תשובה שגויה. ניתן לראות כי עבור $t = log(128/\epsilon^2)$ נקבל כי $H'(r)$ מסכימים על $7/8$ מה־rים בהסתברות $3/4$. כלומר , הגענו לאותן תוצאות כמו בנסיון הראשון (זה מאד מפתיע כי חסכנו המון! הסוד הוא שהשתמשנו באי תלות בזוגות ולא באי תלות מלאה). במצב זה נוכל לשחזר את x בהסתברות גבוהה. בחירת t כנ"ל תבטיח לנו זמן ריצה (וגודל רשימה) פולינומיאליב־$1/\epsilon$. זמן הריצה של האלגוריתם כולו הוא $O(n^2 \cdot \epsilon^{-4} \cdot logn)$. ■

כעת נראה כיצד משפט $G\&L$ נובע מן הלמה האחרונה שהראינו. **הוכחה:** נגדיר אלגוריתם A' המקבל קלט y ומריץ את האלגוריתם מהלמה האחרונה עבור $H(r) = A(y,r)$ (A - האלגוריתם מתנאי משפט $G\&L$ המקורי) אשר מוציא רשימה L. A' בודק כעת עבור כל $x \in L$ אם $f(x) = y$ ומדפיס x כזה אם הוא מוצא. מהנחת המשפט ($G\&L$ המקורי) אנו יודעים כי

$$Pr_{x,r}[A(f(x),r) =< x,r >] \geq \frac{1}{2} + \epsilon$$

ע"י שימוש במרקוב (ווריאציה של אי־ השיוויון כפי שהצגנו לעיל) קל להראות כי

$$Pr_x[Pr_r[A(f(x),r) =< x,r >] \geq \frac{1}{2} + \frac{\epsilon}{2}] \geq \frac{\epsilon}{2}$$

אם כך, אם ניקח באופן מקרי $x \in \{0,1\}^n$ אזי בהסתברות לפחות $\epsilon/2$, x "טוב" עבורנו - קיימת עבורו פונקציה H שעומדת בתנאי הלמה האחרוטנה שהראינו. עבור x "טוב", בהסתברות $1/2$, x יופיע ברשימה L. בסך הכל נקבל כי בהסבתרות לפחות $\epsilon/4$ A' הופך את f כנדרש. ■

משפט 3.40 תהי $f : \{0,1\}^n \to \{0,1\}^n$ פרמוטציה חד־כיוונית, B ביט קשה עבור f. אזי $G : \{0,1\}^n \to \{0,1\}^{n+1}$ המוגדרת כך

$$G(x) = f(x)B(x)$$

מהווה יוצר פ"א.

הוכחה: נניח בשלילה כי D אלגוריתם פולינומיאלי המבחין בין U_{n+1} ל־$G(U_n)$. כלומר

$$|Pr_{x\sim U_n}[D(f(x)B(x))=1]-Pr_{x\sim U_n,b\sim U_1}[D(f(x)b)=1]|>\epsilon \qquad (1)$$

(אנו מסתמכים על כך ש־f פרמוטציה ולכן אם $x\sim U_n$ אזי גם $f(x)\sim U_n$)
תחילה "ניפתר" מהערך המוחלט ב־(1) והאי שיוויון שיתקבל יתקיים עבור D או משלימו (שרץ בסיבוכיות זמן זהה).
נתאר אלגוריתם A שיפעל כדילקמן:
בהינתן קלט $y=f(x)$, A בוחר ביט מקרי $r\sim U_1$. אם $D(yr)=1$ אזי A יפלוט r. אחרת, $1-r$. קל לראות ש־A פולינומיאלי. אנו נטען כי מתקיים:

$$Pr_{x\sim U_n}[A(f(x))=B(x)]>\frac{1}{2}+\epsilon$$

ולכן B אינו ביט קשה בסתירה להנחתנו.
על מנת להדגיש כי A תלוי ב־r, הביט המקרי שנבחר אנו נשתמש בסימון A_r. נשים לב כי

$$Pr_{x,r}[A_r(f(x))=B(x)]=Pr_{x,r}[A_r(f(x))=B(x)|r=B(x)]\cdot Pr[r=B(x)]$$

$$+Pr_{x,r}[A_r(f(x))=B(x)|r\neq B(x)]\cdot Pr[r\neq B(x)]$$

בגלל ש־r ביט מקרי, ברור כי ללא תלות בערך $B(x)$ מתקיים

$$Pr[r=B(x)]=Pr[r\neq B(x)]=\frac{1}{2}$$

ולכן, על פי הגדרת A ההסתברות המדוברת מסתכמת ל־

$$\frac{1}{2}\cdot Pr_{x,r}[D(f(x)r)=1|r=B(x)]+\frac{1}{2}\cdot Pr_{x,r}[D(f(x)r=0|r\neq B(x)]$$

הביטוי הימני שווה ל־ $\frac{1}{2}-\frac{1}{2}\cdot Pr_{x,r}[D(f(x)r)=1|r\neq B(x)]$. אנו נוסיף ונחסיר את הביטוי $\frac{1}{2}\cdot Pr_{x,r}[D(f(x)r)=1|r=B(x)]$ ונקבל

$$\frac{1}{2}+Pr_{x,r}[D(f(x)r)h=1|r=B(x)]-\frac{1}{2}\cdot(Pr[D(f(x)r)=1|r=B(x)]+Pr[D(f(x)r)=1|r\neq B(x)])$$

הביטוי בסוגריים הימניים שווה ל־$Pr[D(f(x)r)]=1$. מהנחתנו המוקדמת על D נסיק כי הביטוי כולו גדול מ־$\frac{1}{2}+\epsilon$ כנדרש. ■

בעזרת משפט $G\&L$ והבנייה האחרונה הראינו כי קיום פרמוטציה חד־כיוונית גורר קיום יוצר פ"א ארוך פולינומיאלית כרצוננו. ניתן להראות ע"י בנייה נוספת ומורכבת יותר כי די בקיום פונקציה חד־כיוונית (כלומר, לא נדרש שתהיה בהכרח פרמוטציה). אנו לא נראה את הבנייה הזאת בקורס. נציין כי דרישה לקיום פונקציה חד־כיוונית היא דרישה די מינימאלית עבור קריפטוגרפיה.

3.5 פונקציות פ"א

נדון כעת בחסרון של יוצרים פ"א (בהקשר הקריפטוגרפי כמובן). בעזרת יוצרים פ"א בוב ואליס יכולים להאריך מפתח קצר למפתח ארוך (פולינומיאלית בגודל המפתח המקורי). מכיוון שבמציאות בוב ואליס שולחים בינהם מספר הודעות בצורה סימטרית, דרוש מנגנון סינכרון שיעזור להם להסכים באיזה חלק של המפתח הם משתמשים ברגע נתון. זה כמובן יותר $overhead$ על ערוץ התקשורת ודורש שמירת מצב מצד בוב ואליס. נתאר מנגנון נוסף הנקרא פונקציות פ"א שיפתור את בוב ואליס מהצרכים הללו. מיד נגדיר את המושג פונקציות פ"א. לפני כן נגדיר את המושג פונקציות מקריות.

הגדרה 3.41 פוקנציה מקרית $F(\cdot)$ הממפה מחרוזות באורך n למחרוזות באורך n היא פונקציה הממפה לכל אחד מ־2^n הקלטים האפשריים מחרוזת אקראית בת n ביטים.

משמעות הדבר היא שנדרשים $2^n \cdot n$ מטבעות על מנת להגדיר פונקציה מקרית. גודל זהה נדרש על מנת לשמור את טבלת האמת של F. נרצה להשיג יכולות דומות במחיר נמוך (פולינומיאלי). אמנם פונקציה המתוארת ע"י n ביטים רחוקה מלהיות פוקנציה מקרית אך בכל זאת אנו נראה שתחת ההנחה (אקסיומת PRG) כי קיים יוצר פ"א ניתן להראות כי קיים אוסף של פונקציות פ"א המוגדר בצורה הבאה:

הגדרה 3.42 יוצר/אוסף של פונקציות פ"א $\mathcal{F} = \{f_s\}_{s\in\{0,1\}^*}$ זה אוסף פונקציות כך שלכל $s \in \{0,1\}^*$, f_s מעתיקה מחרוזות באורך $|s|$ ביטים למחרוזות באורך s ביטים (למעשה, התמונה יכולה להיות בגודל פולינומיאלי ב־s. אנו נעסוק לרב במקרה הסימטרי) המקיימת את התכונות הבאות:

1. f ניתנת לתיאור ולחישוב בזמן פולינומיאלי.
2. נחשוב על שני המשחקים הבאים:

משחק 1

s נבחר אקראית מתוך $\{0,1\}^n$.

יריב מקבל גישה לפונקציה $f_s(\cdot)$ והוא יכול לבצע קריאות לפונקציה כל עוד הוא חפץ (אך לא יותר מ־$poly(n)$ שזה זמן הריצה שלו).

יריב מוציא פלט $v \in \{0,1\}$.

משחק 2

נבחרת פונקציה מקרית $F : \{0,1\}^n \to \{0,1\}^n$.

יריב מקבל גישה לפונקציה $F(\cdot)$ והוא יכול לבצע קריאות לפונקציה כל עוד הוא חפץ (אך לא יותר מ־$poly(n)$ שזה זמן הריצה שלו).

יריב מוציא פלט $v \in \{0,1\}$.

תחת הגדרת המשחקים האלה אנו דורשים שלכל יריב פולינומיאלי המפעיל אלגוריתם A קיימת פונקציה $\epsilon : \mathbb{N} \to [0,1]$ זניחה פולינומיאלית כך שמתקיים

$$|Pr[A\,outputs\,1\,in\,game\,1] - Pr[A\,outputs\,1\,in\,game\,2]| < \epsilon(n)$$

נראה כעת כי אקסיומת PRG גוררת קיום אוסף פונקציות פ"א.

משפט 3.43 אם אקסיומת PRG נכונה (כלומר, קיים יוצר פ"א) אזי קיימות פונקציות פ"א.

הוכחה: (מיכאל לא נתן בכיתה הוכחה מלאה אלא הוכחה "בנפנוף ידיים". אנו נסקור כאן את ההוכחה בצורה מעט יותר מקיפה) נניח כי קיים יוצר פ"א $G : \{0,1\}^n \to \{0,1\}^{2n}$. אנו ניעזר ב־$G$ על מנת לבנות אוסף פונקציות פ"א f_s. נבנה עץ בינארי מלא בגובה n באופן הבא: השורש יקבל את הערך $s \in U_n$ (ה־$seed$ של הפונקציה). לכל קודקוד פנימי עם תיוג v, בנו הימני יקבל את הערך $G_1(v) = G(v)_{[n+1\ldots 2n]}$ ובנו השמאלי יקבל את הערך $G_0(v) = G(v)_{[1\ldots n]}$. יש לנו 2^n עלים. ניתן לזהות כל עלה ע"י מחרוזת ב־$\{0,1\}^n$ לפי השביל מהשורש אל העלה כאשר 0 מבטא בחירה בבן שמאלי, 1 מבטא בחירה בבן ימני. אנו מגדירים את $f_s(x)$ להיות התיוג של העלה המתאים ל־x. באופן פורמאלי

$$f_s(x) = (G_{x_n-1}(\ldots G_{x_1}(s))\ldots)$$

נשים לב כי למרות שגודל העץ הינו אקספוננציאלי ב־n, חישוב $f_s(x)$ דורש חישוב של הפונקציה G (החשיבה בזמן פולינומיאלי) לאורך ענף באורך n.

עלינו להראות כעת כי הבנייה שהצגנו מהווה אוסף פונקציות פ"א. נניח בשלילה כי קיים יריב ואלגוריתם A בסיבוכיות זמן T אשר מבחין בין $f_s(\cdot)$ לפונקציה מקרית בהסתברות גדולה מ־ϵ. אנו נראה כי קיים אלגוריתם A' בעל סיבוכיות זמן T' השקולה פולינומיאלית ל־T, המצליח להבחין בין $G(U_n)$ ל־U_{2n} בהסתברות הגדולה מ־ϵ' השקול פולינומיאלית ל־ϵ.

נניח מספר הנחות על היריב אשר אינן משנות את הסתברות ההצלחה שלו (אך מקלות עלינו בניתוח):

1. הוא מבצע בדיוק T שאילתות (אם הוא מבצע פחות נגיד שהוא משלים ל־T שאילתות ע"י שאילת שאילות חסרות משמעות).
2. הוא לא שואל את אותה שאלה פעמיים (לצורך העניין הוא מחזיק טבלה ובה תשובות לשאלות שכבר שאל והוא יכול להעזר בה במקום לשאול שוב).

כעת נתאר את האינטרקציה של היריב עם $f_s(\cdot)$. תחילה העץ יכיל רק את השורש בעל הערך s. בכל עת שהיריב מבצע שאילתה $f_s(x)$ עבור x כלשהו היריב יחשב את הערך המתאים תוך הישענות על חישובים קודמים: נסמן ב־v את הקודקוד הנמוך ביותר בענף שבין השורש לעלה x אשר חושב כבר. היריב ימשיך את החישוב עד העלה x, ישמור את הערכים (לחישובים הבאים) ויחזיר את הערך הסופי. נשים לב שכאשר אנו מחשבים עבור קודקוד v בעל ערך x את הערכים של בניו $G_0(x)$, $G_1(x)$ ושומרים אותם, אין עוד צורך לשמור את הערך של v. כמו כן נשים לב כי היריב יבצע לכל היותר $M = T \cdot n$ קריאות ל־G.

אנו נשתמש כעת בטיעון היברידי. נסמן עבור $i = 0, ..., m$ את H_i להיות ההשקפה של היריב על העץ שבנינו עבור $f_s(\cdot)$ מלבד שב־i הפעמים הראשונות בהן נדרשת הפעלה של G לתיוג שני בנים של קודקוד v כלשהו בעד ערך x, מבצעים "$fake\,invocation$": במקום לתייג את בניו של v ע"י הערכים $G_0(x), G_1(x)$, מגרילים $x_0, x_1 \in \{0,1\}^n$ ומתייגים את הבנים בעזרת ערכים אלו (ומוחקים את הערך של v). קל לראות כי H_0 זהה להשקפה של היריב על f_s בעוד ש־H_m זהה להשקפה של היריב על פונקציה מקרית. אנו רוצים להראות שאם ניתן להבחין בין H_0 ל־H_m אזי ניתן להבחין בין $G(U_n)$ לבין U_{2n}. מטיעונים היברידיים, אם ניתן להבחין בין H_0 ל־H_m אזי קיים $i \in [m]$ כך שניתן להבחין בין H_i לבין H_{i-1}. נניח אם כן כי קיים אלגוריתם C המבחין בין H_i לבין H_{i-1}. אנו נבנה אלגוריתם C' המבחין בין $G(U_n)$ לבין U_{2n}. בהנתן קלט $y \in \{0,1\}^{2n}$ (כאשר y בא מההתפלגות המקרית U_{2n} או מ־$G(U_n)$), C יריץ את האלגוריתם של היריב (T שאילתות על T מחרוזות שונות תוך כדי בניית העץ כפי שתיארנו קודם) בשינוי קל: כאשר מגיעים לחישוב ה־i , H_{i-1} מסתכל על הערך x בקודקוד הנוכחי v, מחשב $G(x) = (x_0, x_1)$ ומקצה את הערכים x_0, x_1 לבנו הימני והשמאלי בהתאמה (ומוחק את x). H_i לעומתו אמור להגריל x_0, x_1 מקריים ולהקצות לבים של v. C יבחר $(x_0, x_1) = y$ לצורך התיוג. אם $y \sim U_{2n}$ אכן קיבלנו את H_i. אם $y \sim G(U_n)$ קיבלנו את H_{i-1}. C' יציג ל־C את H_{i-1} שבנה. אם C יאמר שמדובר ב־H_i, C' יכריע ש־y בא מ־U_{2n}. אם C יאמר שמדובר ב־H_{i-1}, C' יאמר יכריע ש־y בא מ־$G(U_n)$. ברור כי סיכויי ההצלחה של C' זהים לאלו של C וכי C' פולינומיאלית "איטי" יותר מ־C. ■

$Security\,Against\,Chosen - Plaintext\,Attacs\,(CPA)$ **3.6**

עד כה הצגנו מודלים בהם איב הינה פאסיבית - לומדת בעיקר ע"י האזנה לערוץ בו אליס ובוב מדברים (גם במודלים שהצגנו באמצעות משחקים, איב לכל היותר בוחרת מספר הודעות מצומצם ומבקשת את ההצפנות שלהם). כעת, נדבר על מודל התקפה הנקרא $chosen - plaintext\,attack\,(CPA)$. מיד לאחר מכן נגדיר אבטחה כנגד סוג זה של התקפה. ההגדרה תהיה דומה להגדרות קודמות מלבד "הכח" הנוסף שינתן לאיב.

השוני העיקרי הוא שכעת איב יכולה לבקש מספר הצפנות רב כרצונה כשמגבלותיה היחידות נובעות מהגבלת זמן הריצה (לרב אנו מגבילים את איב לרוץ בזמן פולינומיאלי). איב למעשה מקבלת גישה ל"קופסא שחורה" המצפינה הודעות לפי בחירתה. עם זאת, איב לא מכירה את המפתח k. דרישתנו מסכימת הצפנה הבטוחה מפני CCA היא שאיב לא תוכל להבחין בין שתי הצפנות של שתי הודעות שונות. נתאר זאת בצורה מעט יותר פורמאלית.

הגדרה 3.44 סכימת הצפנה בטוחה מפני CPA

נתאר את כללי המשחק:

1. אליס בוחרת מפתח $k \in U_n$.
2. איב בוחרת שתי הודעות $m_0, m_1 \in \{0,1\}^n$ ושולחת אותם לאליס.
3. אליס בוחרת $i \leftarrow_R \{1,2\}$ ושולחת את $y = E_k(m_i)$ לאיב.
4. תחת המגבלה שזמן הריצה שלה חייב להיות $poly(n)$, איב רשאית לדרוש הצפנות להודעות שונות. בין היתר, איב רשאית לדרוש הצפנות עבור ההודעות m_0, m_1.

5. על סמך האינטרקציה הממושכת, איב מנסה להסיק את זהות i. כלומר, איב בוחרת $j \in \{1,2\}$.
6. איב מנצחת במשחק אם $j = i$.

סכימת הצפנה (E, D) תקרא $CPA-secure$ אם קיימת פונקציה זניחה $\epsilon(\cdot)$כך שסיכויי ההצלחה של איב במשחק הנ"ל הם לכל היותר $\frac{1}{2} + \epsilon(n)$.

הערה 3.45 נשים לב כי סכימת הצפנה דרטרמיניסטית אינה יכולה להיות $CPA - secure$. זאת מכיוון שאיב יכולה לדרוש במפורש את ההצפנות של m_0, m_1 ולהשוואנו נראה שההסתברות ל"מאורע הרע" (חזרה על r_c במהלך ריצת האלגוריתם) קורה בהסתברות נמוכה ומכאן נסיק את ל־y ומכך להסיק בוודאות את זהות i. כמו כן, נשים לב כי סכימת הצפנה $CPA - secure$הינה בטוחה חישובית שכן ב־CCA רק הוספנו "כח" לאיב (ניתן לחשוב על איב המאזינה כאיב היוזמת CCA אך בוחרת שלא לנצל את הגישה שלה ל"קופסא השחורה" $E_k(\cdot)$.

נתאר סכימת הצפנה העושה שימוש בפונקציות פ"א. אנו נראה כי הסכימה הזו הינה $CPA - \text{secure}$.

סכימת הצפנה $CPA - \textbf{secure}$ נתאר כעת סכימת הצפנה מקובלת המשתמשת באוסף פונקציות פ"א.

- בוב ואליס מכירים אוסף פונקציות פ"א $\mathcal{F} = \{f_s\}_{s \in \{0,1\}^*}$.
- נבחר מפתח $k \in U_n$.
- בוחרים $r \in U_{\mathrm{n}}$. בהנתן הודעה $x \in \{0,1\}^n$ אליס שולחת את $E_k(x) = (r, f_k(r) \oplus x)$.
- בהנתן (r, y) בוב מפעיל $f_k(r) \oplus y$ על מנת לשחזר את x.

(נשים לב כי $D_k(E_k(x)) = f_k(r) \oplus f_k(r) \oplus x = x$ כנדרש)

באופן אינטואיטיבי, ניתן להשתכנע כי כל עוד הערך r באמצעותו מוצפנת ההודעה m_i (כאשר $i \in \{0,1\}$ ו־ m_0, m_1 הן ההודעות המקוריות שבחרה איב) אשר נסמנו ב־r_c. אינו חוזר על עצמו במהלך המשחק, $f_k(r_c)$ נראה מקרי לחלוטין עבור איב. מכאן, גם $f_k(r) \oplus m_i$ נראה מקרי לכל הודעה. מכאן, שהסכימה הזאת דומה לסכימת $one - time\, pad$ שהכרנו.

משפט 3.46 סכימת ההצפנה הנ"ל הינה $CPA - secure$.

הוכחה: רעיון ההוכחה: אנו נראה כי המערכת הינה בטוחה לו היינו עושים שימוש בפונקציה מקרית ולא ב־ $\mathcal{F}$ אשר הינה פ"א. לאחר מכן, נראה שאם המערכת אינה בטוחה תחת שימוש ב־$\mathcal{F}$ אז נוכל להבחין בין פונקציה מקרית לבין $\mathcal{F}$ בהסתברות לא זניחה וזאת בסתירה לכך ש־$\mathcal{F}$ הינה פ"א.

נזכור כי זמן הריצה של איב הינו פולינומיאלי. נניח כי איב מבצעת $q(n)$ שימושים ב"קופסא השחורה" שעומדת לרשותה.

נניח כי ברשותנו פונקציה מקרית $F(\cdot)$ונחשוב על שימוש בסכימת ההצפנה שתיארנו כאשר אנו מחליפים את $\mathcal{F}$ ב־F. נסמן ב־r_c כמקודם את המחרוזת האקראית אשר בעזרתה מוצפנת ההודעה הראשונה (m_0 או m_1 בהתאם לביט i). נשים לב כי הדרך היחידה של איב לקבל אינפורמציה על i היא שבמהלך הגישה שלה ל"קופסא השחורה" יעשה שימוש נוסף במחרוזת r_c (אחרת, איב לא לומדת דבר על הערך $F_k(r_c)$). אם באחת הגישות נעשה שימוש נוסף ב־r_c על מנת להצפין הודעה נוספת m איב יכולה לחלץ את $F_k(r_c)$ (איב מקבלת $< r, s >$ וע"י הפעלת $\mathrm{m} \oplus \mathrm{s}$ תוכל לקבל את $F_k(r_c)$). כעת, $F_k(r_c) \oplus y$ יספק לאיב את זהות m_i. הסיכוי שבאחת מ־$q(n)$ הגישות נעשה שימוש נוסף ב־r_c על מנת להצפין הודעה נוספת חסום ע"י $\frac{q(n)}{2^n}$. אם אין חזרה על הערך r_c, הערך $F_k(r_c)$ נראה לאיב מקרי לחלוטין ולכן סיכויי ההצלחה שלה במצב זה הם בדיוק $\frac{1}{2}$. מחסם האיחוד נסיק כי סיכויי ההצלחה של איב במשחק עבור F מקרית הם לכל היותר $\frac{1}{2} + \frac{q(n)}{2^n}$.

כעת, נניח כי סיכויי ההצלחה של איב במשחק עבור $\mathcal{F}$ פ"א הינם $\frac{1}{2} + \epsilon(n)$. אנו רוצים להראות כי $\epsilon(n)$ זניחה. נניח בשלילה כי $\epsilon(n)$ אינה זניחה. אזי, ההפרש בין סיכויי ההצלחה של איב במשחק עם פונקציה פ"א לסיכויי ההצלחה במשחק עם פונקציה מקרית אינו זניח (מתקבל ההפרש $\epsilon(n) - \frac{q(n)}{2^n}$. מכיוון ש־$q(n)$ הוא פולינום, ברור כי $\frac{q(n)}{2^n}$ זניחה. כמו כן, סכום פונקציות זניחות הוא פונקציה זניחה). אם כך, ניתן להשתמש באלגוריתם A שמפעילה איב על מנת לבנות אלגוריתם D המבחין בין $\mathcal{F}$ ל־F בהסתברות לא זניחה: בהנתן פונקציה $g : \{0,1\}^n \to \{0,1\}^n$,D יסמלץ את התפקיד של אליס מול האלגוריתם A (למעשה ישחק את שני התפקידים בהתקפת CPA). אם איב "מנצחת" במשחק, D יודיע כי g פ"א. אחרת, יודיע כי g הינה מקרית (על מנת להתאים את עצמנו למודל, נאמר ש־D יודיע "1" אם"ם איב מנצחת במשחק). נקרא למשחק עם $\mathcal{F}$ "משחק 1" ולמשחק עם F "משחק 2". קל לראות כי מתקיים.

$$|Pr[A\,outputs\,1\,in\,game\,1] - Pr[A\,outputs\,1\,in\,game\,2]| < \epsilon(n) - \frac{q(n)}{2^n}$$

כאמור ההפרש $\epsilon(n) - \frac{q(n)}{2^n}$ אינו זניח. מכאן, D מבחין בין Fלבין $\mathcal{F}$ בסתירה לכך ש־$\mathcal{F}$ פ"א. ■

3.7 $Security\,Against\,Chosen - Chipertext\,Attacs\,(CCA)$

נמשיך ונוסיף "כח" לאיב. עד כה, עסקנו ב־2 סוגי יריבים: יריב המסוגל להאזין בצורה פאסיבית ויריב שבאופן אקטיבי בוחר הודעות ומבקש הצפנות שלהן (CPA). כעת נתאר מודל התקפה הנקרא $Chosen-C$hipertext $Attacs\,(CCA)$. במודל זה, מלבד היכולות שניתנו לאיב במודל CPA, אנו מאפשרים לה גם גישה ל"קופסא שחורה" נוספת באמצעותה היא יכולה לפענח הודעות מוצפנות (תחת מגבלה אחת שמיד נדון בה). למעשה, כללי המשחק דומים לכללי המשחק במודל CPA. נתאר בצורה פורמאלית מהי סכימת הצפנה הבטוחה מפני CCA.

הגדרה 3.47 סכימת הצפנה בטוחה מפני CCA
נתאר את כללי המשחק:
1. אליס בוחרת מפתח $k \in U_n$.
2. איב בוחרת שתי הודעות $m_0, m_1 \in \{0,1\}^n$ ושולחת אותם לאליס.
3. אליס בוחרת $i \leftarrow_R \{1,2\}$ ושולחת את $y = E_k(m_i)$ לאיב.
4. תחת המגבלה שזמן הריצה שלה חייב להיות $poly(n)$, איב רשאית לדרוש הצפנות להודעות שונות ופיענוח הודעות שונות. בין היתר, איב רשאית לדרוש הצפנות עבור ההודעות m_0, m_1. איב אינה רשאית לדרוש פיענוח של y.
5. על סמך האינטרקציה הממושכת, איב מנסה להסיק את זהות i. כלומר, איב בוחרת $j \in \{1,2\}$.
6. איב מנצחת במשחק אם $j = i$.

סכימת הצפנה (E,D) תקרא $CCA-secure$ אם קיימת פונקציה זניחה $\epsilon(\cdot)$כך שסיכויי ההצלחה של איב במשחק הנ"ל הם לכל היותר $\frac{1}{2} + \epsilon(n)$.

הערה 3.48 נשים לב כל מבלי להגביל את איב כך שלא תוכל לדרוש פיענוח של y, אין ביכולתנו לבנות סכימת הצפנה בטוחה עבור המודל הנ"ל.

אנו לא נראה בשלב זה סכימת הצפנה $CCA - secure$. על מנת להראות כמה כח נוסף לאיב במודל הנ"ל נראה כי אותה סכימה לגביה הראינו כי היא $CPA - secure$ אינה $CCA - secure$.

משפט 3.49 סכימת ההצפנה אשר הראינו למעלה אינה $CCA - secure$.

הוכחה: נתאר התקפה של איב: איב בוחרת $m_0 = 0^n, m_1 = 1^n$. אליס בוחרת $i \leftarrow_R \{0,1\}$ ושולחת לאיב את $E_k(m_i) =< r, F_k(r) \oplus m_i >$. איב, מצידה מקבלת את $c =< r, s >$, אינה יכולה לדרוש את הפיענוח של $< r, s >$ אך יכולה להפוך את הביט הראשון של s, לקבל את המחרוזת s' ולדרוש את הפיענוח של $c' =< r, s' >$. אם הפיענוח המתקבל הינו 10^{n-1} איב תדע כי i=0. אחרת, הפיענוח המתקבל הינו 01^{n-1} ואז איב תדע כי $i = 1$. ■

4 $Secure\,2 - party\,and\,multiparty\,computation$

לעיתים תכופות עולה הצורך לבצע חישוב על סמך נתונים המצויים בידי גורמים שונים כך שבסוף התהליך כל משתמש יחשף לתוצאת החישוב אך לא ילמד פרט נוסף אודות הנתונים המצויים בידי הגורמים האחרים. לפני שנגדיר פורמאלית את הבעיה, נציג בעיה מוכרת.

בעית המיליונרים של Yao A, B מחזיקים כל אחד סכום כסף. הם מעוניינים לחשב למי מבינהם סכום כסף גדול יותר מבלי שכל צד ילמד מהו הסכום שמחזיק השני. נניח כי הסכום שמחזיק A מיוצג ע"י מחרוזת $x \in \{0,1\}^n$ והסכום שמחזיק B מיוצג גם כן ע"י מחרוזת $y \in \{0,1\}^n$. נציע שתי פונקציות לחישוב:

1. $f(x,y)$ תפלוט 1 אם $x > y$ ותפלוט 0 אם $y \leq x$. קל לראות שאם A, B מעבירים בצורה חשאית את הקלט שלהם לצד שלישי אשר מחשב את $f(x,y)$ ומחזיר להם את התשובה, אין הם יכולים ללמוד את הסכום שמחזיק הצד השני.

2. $f(x,y)$ תפלוט $2x$ אם $x > y$. אחרת, תפלוט $2y+1$. נשים לב כי אחת הפלטים הוא בהכרח זוגי והשני בהכרח אי זוגי. קל לראות כי אם A, B מעבירים בצורה חשאית את הקלט שלהם לצד שלישי אשר מחשב את $f(x,y)$ ומחזיר להם את התשובה, אזי בעל הסכום הנמוך מבין השניים יכול ללמוד מהו הסכום שמחזיק בעל הסכום הגבוה (למשל אם מתקבל הפלט $2x$, שני הצדדים רואים כי הסכום זוגי ומסיקים כי ל־A סכום גבוה יותר. B מחלק את הפלט ב־2 ומסיק כי זה הסכום שמחזיק A.

אנחנו נתעניין במיוחד במצבים בהם לא קיימת האפשרות למסור את הנתונים לצד שלישי, אמין ומוסכם שיבצע את החישוב ויחזיר את התשובה למשתתפים. במצבים שכאלה ננסה למצוא פרוטוקולים לחישוב משותף כך שבסוף התהליך המשתתפים ישיגו בדיוק את המידע שהיו משיגים אם היו נעזרים בצד שלישי. נגדיר את הדברים בצורה מסודרת.

הגדרה 4.1 מספר שחקנים $p_1, ..., p_n$ מחזיקים n ערכים $x_1, ..., x_n$ כך ששחקן p_i מכיר בערך x_i. פרוטוקול לחישוב פונ' $f : x_1 \times x_2 \times ... \times x_n \to y$ יקרא פרטי אם כאשר השחקנים פועלים לפי הפרוטוקול הם לא מקבלים אינפורמציה שאינה נובעת מהקלט שלהם ותוצאת חישוב הפונקציה.

הגדרה 4.2 הפרוטוקול האידיאלי: מוסיפים צד שלישי אמין T. כל שחקן מוסר את x_i ל־T. T מחשב את f ומודיע את התוצאה לכל השחקנים (או לקבוצה חלקית) וגם מודיע מי השתתף (רלוונטי כאשר השחקנים יכולים לסטות מהפרוטוקול).

הגדרה 4.3 נאמר **שפרוטוקול** לחישוב פונקציה הוא t**־פרטי** אם לכל קבוצה S כך ש־$|S| \leq t$ לא מצליחה להשיג מידע נוסף מהפרוטוקול. כלומר, לכל קבוצה S כזו קיים סימולטור (יעיל) אשר לאחר אינטרקציה של השחקנים עם הגורם האמין T במחשק האידיאלי יכול להשלים את כל ההודעות ש־S מקבלת ושולחת במהלך הפרוטוקול (ע"פ ההתפלגות בפרוטוקול).

כלומר, נתבונן במשחק הבא:

1. השחקנים מעבירים את הקלט שלהם לצד שלישי אמין T. T מחשב ומודיע להם את התשובה.
2. השחקנים מריצים את הפרוטוקול. כל צד שומר את המידע שהוא רואה. נסמן את "ההשקפה" של שחקן p_i במהלך הפרוטוקול ($View_{p_i}[(p_1, p_2, ..., p_n)[x_1, ..., x_n]]$.

הפרוטוקול יחשב פרטי אם p_i, בהנתן אינטרקציה של המשחק עם T, תוכל לחשב את $View_{p_i}[(p_1, p_2, ..., p_n)[x_1, ..., x_n]]$.

טענה 4.4 לפונקציה And לא קיים פרוטוקול פרטי.

הוכחה: נניח כי ל־B יש ערך 0. בהנתן תוצאת החישוב $A \, and \, B$, ל־B לא נוספת אינפורמציה על זהות הערך שמחזיקה A (מובטח שתוצאת החישוב תהיה 0 ללא תלות בערך שמחזיקה A). לכן, על מנת שהפרוטוקול יהיה פרטי ההודעה ש־A מוסר ל־B לא יכולה להיות תלויה בערך שמחזיק A (כי אסור שהפרוטוקול ימסור ל־B מידע שלא נובע מהקלט שלו ומתוצאת חישוב הפונקציה). באופן סימטרי, ההודעה ש־B מוסר ל־A לא תלויה בערך שמחזיק B (וכך הלאה...). בצורה זו, ברור כי לא מובטח כי A ו־B יצליחו לחשב את הפונקציה (נובע מהעובדה שהפונקציה And אינה קבועה). ולכן הפרוטוקול לא מקיים את הדרישה הבסיסית - חישוב הפונקציה f. ■

הערה 4.5 ניתן להראות שלכל פונקציה שמכילה בטבלת האמת שלה את הערכים

	y_0	y_1
x_0	0	0
x_1	0	1

אין פרוטוקול פרטי. מסתבר שהתנאי הזה הוא גם תנאי מספיק. כלומר, אם אין ריבוע כזה בטבלת האמת של f, אזי f ניתנת לחישוב פרטי.

נציג כעת בעיה נוספת. נניח כי n מרצים, נסמנם $p_1, ..., p_n$, מתכנסים בקפיטריה ומעוניינים לחשב את ממוצע משכורותיהם. כל שחקן p_i יודע כמובן מהו גובה השכר שלו אך לא יודע את גובה השכר של עמיתיו. נציג פרוטוקול לחישוב הממוצע שבאמצעותו לא נוסף לאף מרצה מידע על משכורתו של מרצה אחר (מובן שניתן להשיג את המטרה הזאת באמצעות צד שלישי).

אנו נזדקק להנחה הבאה: לכל $i \in [n]$, מתקיים $x_i \leq B$. אם כך, ברור כי $\Sigma_{i=1}^{n} x_i \leq n \cdot B$. נסמן $M = n \cdot B$. נשים לב כי $\Sigma_{i=1}^{n} x_i = \Sigma_{i=1}^{n} x_i \, (mod \, M)$. מעתה, כל פעולות האריתמטיקה שנתאר יהיו מודלול M. כעת, נתאר את הפרוטוקול:

1. p_1 בוחר r מקרי בתחום $[0, M]$.

2. p_1 מחשב $r + x_1 \,(mod\, M)$ ומוסר לשחקן p_2.

3. p_2 מוסיף לסכום שקיבל את x_2 (גם כן, מודולו 2) ומעביר ל־p_3 וכך הלאה עד ש־p_n מוסיף לסכום שקיבל את x_n ומוסר ל־x_1.

4. p_1 (היחיד שמכיר בערך r) מחסיר את r מהסכום שקיבל, מחלק בn ומשתף את כולם בתוצאה.

ברור כי הפרוטוקול מחשב את הפונקציה. על מנת להשתכנע כי הוא פרטי, ניזכר כי r נדגם באופן מקרי ולכן ניתן להראות באינדוקציה שהערך שרואה כל שחקן הוא מקרי ולכן אין ביכולתו להסיק מידע נוסף.
אנו אומרים שהפרוטוקול שהראינו הוא 1־פרטי. נראה שהפרוטוקול אינו 2־פרטי. כלומר, נראה ששני שחקנים יכולים להסיק ביחד מידע שלא היו מסיקים מגורם שלישי. למשל, השחקנים p_3, p_5 יכולים להסיק יחדיו מהו השכר של השחקן p_4 ע"י החסרת הערך שמוסר p_3 ל־p_4 מהערך שמקבל p_5 מ־p_4. נראה כעת פרוטוקל n־פרטי לחישוב ממוצע השכר של המרצים:

1. כל שחקן p_i מגריל $n-1$ ערכים מקריים בתחום $[0, M]$ $r_{i,1}, ..., r_{i,n-1}$. כמו כן, הוא קובע את הערך $r_{i,n}$ כך ש־$\Sigma_{j=1}^{n} r_{i,j} = x_i \,(mod\, M)$.

2. כל p_i מוסר בבטחה את $r_{i,j}$ ל־p_j לכל j.

3. כל p_j מקבל את $r_{i,j}$ לכל i (נובע מ־2) ומחשב את $S_j = \Sigma_{i=1}^{n} r_{i,j} \,(mod\, M)$.

4. כולם מחברים את הסכום $\Sigma_{j=1}^{n} S_j \,(mod\, M)$ ומקבלים את הסכום. נותר רק לחלק ב־n על מנת לקבל את הממוצע.

קל להשתכנע כי לכל קבוצה של שחקנים, אין ביכולתם להסיק מידע שלא היו מסיקים מגורם שלישי.

משפט 4.6 $p_1, ..., p_n$ שחקנים. לכל פונקציה f יש פרוטוקול t־פרטי כאשר $n > 2t$. הפרוטוקול פולינומיאלי במספר השחקנים ובגודל המעגל לחישוב f.

הוכחה: נוכיח עבור $n = 2t + 1$ וכמובן שזה יהיה תקף עבור n גדול יותר. עלינו להראות שכל קבוצה בגודל t (או פחות) לא תקבל אינפורמציה נוספת. נראה עוד כי כל קבוצה בגודל $t + 1$ תדע את הסוד (זה כמובן לא מחוייב אבל נוח לצורכי ההוכחה). נתאר את השלבים השונים במהלך הפרוטוקול. **בשלב הקלט** כל שחקן מכניס את הקלטים שלו למערכת בעזרת חלוקת סוד. בוחרים מספר ראשוני p כך ש־$p > n$. אנו נעבוד בשדה F_p. נקבעים באופן משותף ומקרי ע"י כל השחקנים n איברים שונים מאפס ושונים זה מזה $\alpha_1, ..., \alpha_n$בשדה. נתאר חלוקת סוד של שחקן p_i. השחקן בוחר t איברים מקריים $a_1, ..., a_t$ בהתפלגות אחידה בשדה זה ויוצר את הפולינום $s_i(x) = x_i + a_1 \cdot x + ... + a_t \cdot x_t$. כעת השחקן מעריך את הפולינום ב־n הנקודות ומוסר לשחקן p_j את הערך $s_j = s(\alpha_j)$.
לפני שנתאר את שלב החישוב, נשים לב כי כל קבוצה בגודל $t + 1$ יכולה ע"י חישוב פולינום אינטרפולציה למצוא את כל מקדמי $s_i(x)$ לכל $i \in [n]$ ומכאן לחשב את $s_i(0)$ וכך לגלות את "הסוד" של p_i. נרצה להראות שכל קבוצה בגודל לכל היותר t לא מקבלת אינפורמציה על הסוד. נרצה להראות טענה שתסייע לנו בהמשך:
טענה: $s_1, ..., s_t$ מפולגים בהתפלגות אחידה על פני F_p^t.
הוכחה: כאמור, $\alpha_1, ..., \alpha_n$ ידועים וקבועים. נשים לב כי בהנתן המקדמים $a_1, ..., a_t$ ניתן להגדיר העתקה שמקבלת את המקדם החופשי s (הסוד) ומוצאת את הערכים $s_1, ..., s_t$. נשים לב כי ההעתקה הינה חח"ע. שכן, בהנתן $s_1, ..., s_t$ ובהנתן s (הערך ב־0) ניתן לקבוע את זהות $a_1, ..., a_t$ ע"י חישוב פולינום אינטרפולציה. מכיוון ש־$a_1, ..., a_t$ מקריים ונדגמים בהתפלגות אחידה והעתקה שהגדרנו הינה חח"ע, נסיק כי גם $s_1, ..., s_t$ מפולגים בהתפלגות אחידה על פני F_p^t כנדרש.
נפנה לתאר את שלב החישוב. **בשלב החישוב** אנו נניח כי המעגל שמחשב את הפונקציה f נתון ע"י מעגל אריתמטי מעל השדה (שערי חיבור וכפל). נניח שנתון **שער חיבור**. נניח כי ערכי הכניסה לשער החיבור a, b כבר חושבו. כלומר, השחקנים מחזיקים בצורת חלוקת סוד את הפולינומים $A(x), B(x)$ אשר מקודדים בהתאמה את הסודות a, b (כך ש־$A(0) = a, B(0) = b$). נסמן $C(x) = A(x) + B(x)$. אם כן, כל שחקן p_k יכול לחשב את $C(\alpha_k) = A(\alpha_k) + B(\alpha_k)$. הערך $c = a + b$ מתקבל ע"י $C(0)$. אם כן, בפעולת החיבור הנחנו כי השחקנים מכירים בצורת חלוקת סוד את הערכים a, b והראינו כיצד מתבצעת פעולת חיבור כך שבסופה השחקנים חולקים את הסוד ־ הערכים של הפולינום C המקודד את הערך c. באופן דומה נבצע הכפלה בקבוע c של הערך a (אנו מניחים שהשחקנים מכירים בצורת חלוקת סוד את הערך a ע"י הפולינום $A(x)$. כל שחקן p_k יחשב את $c \cdot A(\alpha_k)$. השחקנים יכירו אם כן את ערכי הפולינום $c \cdot A(x)$ וע"י כך יחזיקו את הערך $c \cdot a$ בצורת חלוקת סוד.

כעת נניח שנתון שער כפל ושוב נניח כי ערכי הכניסה לשער החיבור a, b כבר חושבו והשחקנים מחזיקים בצורת חלוקת סוד את הפולינומים $A(x), B(x)$ אשר מקודדים בהתאמה את הסודות a, b. נשים לב כי פולינום המכפלה $D(x) = A(x) \cdot B(x)$ אינו פולינום מקרי מדרגה $2t$. נניח למשל כי הפולינומים A, B מדרגה 1. אזי אם $a = 0, b = 0$ המקדם של x בפולינום המכפלה הינו 0. אחרת, בסבירות גבוהה המקדם אינו 0. מכיוון שלא היינו רוצים שזהות הפולינום $C(x)$ "תסגיר" את הסודות a, b נאלץ לבצע כפל בצורה מעט יותר מורכבת. כל שחקן p_k יכול לחשב את $D(\alpha_k) = A(\alpha_k) \cdot B(\alpha_k)$. הפולינום D הוא מדרגה $2t \leq n-1$ (לכל היותר). אנו נניח כי דרגתו היא $n-1$ (תמיד אפשר לאפס מקדמים). כעת, השחקן p_k מכניס את הערך $D(\alpha_k)$ לחישוב בעזרת חלוקת סוד בעזרת הפולינום ממעלה t $E_k(x)$ (כלומר שוב בוחר ערכים מקריים $a_1, ..., a_t$ בהתפלגות אחידה בשדה שביחד עם המקדם החופשי $D(\alpha_k)$ מהווים מקדמיו של הפולינום. לסיום, הוא מוסר את הערך $E_k(j)$ לשחקן ה־j לכל $j \in [n]$). נסמן $e_k = E_k(0) = D(\alpha_k)$. נטען כי קיימים קבועים $q_1, ..., q_n$ כך ש־$D(0) = \Sigma_{i=1}^n q_i \cdot e_i$. נסמן $C(x) = \Sigma_{i=1}^n q_i \cdot E_i(x)$ ואז $C(0) = D(0) = a \cdot b$. נוכיח כי טענתנו האחרונה נכונה:

נזכור כי הפולינום $D(x) = d_0 + d_1 \cdot x + ... + d_{n-1} \cdot x^{n-1}$. אם ניקח $(d_0, ..., d_{n-1})$ כך ש־

$$(d_0, ..., d_{n-1}) \cdot \begin{pmatrix} 1 & ... & 1 \\ \alpha_1 & & \alpha_n \\ .. & & .. \\ .. & & .. \\ \alpha_1^{n-1} & & \alpha_n^{n-1} \end{pmatrix} = (D(\alpha_1), ..., D(\alpha_n)) = (e_1, ..., e_n)$$

כידוע, מטריצת ונדרמונדה הפיכה ולכן אם נסמן ב־v את מטריצת ונדרמונדה שלהלן נקבל:

$$(e_1, ..., e_n) \cdot v^{-1} = (d_0, ..., d_{n-1})$$

קל לראות כי העמודה הראשונה של v^{-1} היא $(q_1, ..., q_n)^T$. נשים לב עוד כי כל אחד מהשחקנים יכול לחשב את מטריצת ונדרמונדה, את ההפכית שלה ובפרט לדעת את העמודה $(q_1, ..., q_n)^T$. כעת, כל שחקן p_k יודע את הערך $C(\alpha_k)$. אם כך, במקום ליצור חלוקת סוד של ערכי הפולינום $D(x)$ שהינו פולינום לא מקרי מדרגה $2t$, יצרנו חלוקת סוד של הפולינום $C(x)$ שהינו פולינום מקרי מדרגה t שמקדמו החופשי/ערכו ב־0 הינו הערך המבוקש $a \cdot b$.

בשלב הפלט, כל שער שחושב בעזרת קידוד שאת התוצאה שלו צריך לקבל שחקן מסויים ־ כל השחקנים שולחים אליו את חלקי הסוד שלהם והוא מקבל את התוצאה.

נשלים כעת הוכחתנו ונראה כי קבוצה בגודל t אינה לומדת אינפורמציה נוספת. למען נוחות ההוכחה נניח כי מתבצעת הכפלה ב־1 (הפולינום הקבוע 1) לפני כל גילוי פלט. נתחיל מהמקרה המיוחד ־ f מחשבת סכום של הקלטים, כלומר $y = \Sigma_{i=1}^n x_i$. נטען שקבוצה S כך ש־$|S| = t$ אינה לומדת אינפורמציה נוספת. עלינו להראות כי ע"י אינטרקציה עם הפרוטוקול האידיאלי במהלכה שחקני S (ב.ה.כ $S = \{p_1, ..., p_t\}$) מוסרים את x_i ל־T ו־T מחזיר להם את y (אם אחד מחברי S אמורים לקבל את התשובה), הקבוצה S תוכל לחשב את ההודעות שהיתה מקבלת ושולחת במהלך הפרוטוקול המקורי. עבור $p_1, ..., p_t$ נבחר לכל p_k פולינום מקרי s_k המקודד את x_k, ונשלח לכולם את חלקי הסוד. נסמן $y(x) = s_1(x) + s_2(x) + ... + s_n(x)$. נשים לב כי $y = s_1(0) + s_2(0) + ... + s_n(0)$. כל p_k מחבר את n המספרים שהוא קיבל ונסמן את התוצאה ב־r_k. בהינתן $r_1, ..., r_t$ והתוצאה y ניתן לחשב אץ פולינו האינטרפולציה $y(x)$. כדי לשחזר את ההודעות עבור $t+1 \leq k \leq n$ נחשב $y(\alpha_k) = r_k$.

עבור מעגל כללי, קבוצה S בגודל t יכולה לצורך הסימולציה לחלק את הקלטים שלה, לבחור באופן אקראי קלטים עבור שחקנים אחרים ולבנות את ההודעות שמתקבלות מהם. קל לראות שהקבוצה S מייצרת ומקבלת הודעות בדיוק לפי ההתפלגות שנוצרת במהלך הפרוטוקל האמיתי. לפני כל שער שמגלים מתבצע חיבור של n קלטים ואת זה כבר הראינו כיצד אפשר לסמלץ כאשר יודעים את התוצאה. ■

כעת נדבר על **חישוב בטוח**. לעיתים נרצה הגנה גם במקרה שמספר שחקנים סוטים מן הפרוטוקול.

נניח שהצלחנו לחשב קידוד של ערך שחושב. כלומר, יש פולינום $s(x) = s + a_1 \cdot x + ... + a_t \cdot x_t$. כל שחקן p_k מחזיק את $s_k = s(\alpha_k)$. כל שחקן p_k מגלה את $\bar{s_k}$. עבור שחקנים טובים $\bar{s_k} = s_k$ אבל אנחנו מניחים ש־t ערכים יכולים להיות שגויים.

טענה 4.7 אם $n \geq 3t+1$ (אחוז הרעים קטן משליש) אז קיים פולינום יחיד ממעלה לכל היותר t העובר דרך $n-t$ מהנקודות.

הוכחה: נניח ש־$f(x)$ ו־$g(x)$ שני פולינומים ממעלה לכל היותר t וכל אחד מהם עובר דרך $n-t$ מהנקודות. עדיין יש להם לפחות $t+1 \leq n-2t$ נקודות משותפות ולכן כפולינומים ממעלה לכל היותר t מתקיים $f(x) \equiv g(x)$.
הבעיה היא שיש לנו $\binom{n}{t}$ אפשרויות לבחור רעים. לאתר אותם בדרך הטריוויאלית ־ עלות אקספוננציאלית. אבל יש אלג' יעיל:
מתקיים:

$$(a_1, ..., a_t, 0..., 0) \cdot \begin{pmatrix} 1 & ... & 1 \\ \alpha_1 & & \alpha_n \\ .. & & .. \\ .. & & .. \\ \alpha_1^{n-1} & & \alpha_n^{n-1} \end{pmatrix} = (f(\alpha_1), f(\alpha_2), ..., f(\alpha_n))$$

שוב, נסמן ב־v את מטריצת ונדרמונדה לעיל. במידה ויש "סטיות" מתקיים:

$$(f(\alpha_1), ..., f(\alpha_n)) \cdot v^{-1} = (a_1, ..., a_t, 0, ..., 0) + (e_1, ..., e_n)$$

כאשר לכל היותר t מה־e_i אינם אפס. אם ניתן להסתכל על $2t$ האחרים האחרונים בחישוב $(\bar{S_1}, ..., \bar{S_n}) \cdot v^{-1}$ (אפשר פשוט לקחת רק את 2 העמודות האחרונות של v^{-1}) ולקבל את את $(b_1, ..., b_{2t})$. אם $t < n/3$ יש אלג' מתקן שגיאות ל־$b_1, ..., b_{2t}$. ■

נתאר כעת פרוטוקל שידמה את האינטרקציה עם צד שלישי אמין כמקודם תחת ההנחה שלכל היותר $t < n/3$ שחקנים יכולים לסטות מן הפרוטוקול. **בשלב הקלט** ניעזר שוב בחלוקת סוד. D, המחלק (אחד השחקנים או שחקן נוסף) מחזיק סוד S, בוחר פולינום מקרי $f(x) = s + a_1 \cdot x + ... + a_t \cdot x^t$ ($a_i \in F_p$). לאחר מכן, מחלק לכל שחקן p_k את $s_k = f(\alpha_k)$. הנחתנו היא כאמור כי עד t מהשחקנים רעים וגם D יכול להיות רע. רוצים תמיד לקבל את החלוקה כש־D טוב. כש־D רע או שהוא יתאפס או שבהסתברות גבוהה הערכים שאנמאים בידי השחקנים הטובים \הם על פולינום ממעלה $\geq t$.
לצורך הבדיקה: D מחלק עוד $2r$ סודות מקריים בעזרת פולינומים $g_1, ..., g_{2r}$ ממעלה $\geq t$. לכל אינדקס l כאשר $l = 1, ..., r$ נתאים שחקן ל־l מתאים $[p_\ell + l \,(mod\, n)]$. השחן המתאים מפרסם ביט מקרי c_l ומודיע לכולם ול־D מהו הביט. D מפרסם את מקדמי הפולינום $h_l(x) = g_l(x) + c_l \cdot f(x)$ (נשים לב שמכיוון ש־g_l פולינום מקרי כך גם h_l). כל שחקן יכול לבדוק האם $h_l(\alpha_k) = g_l(\alpha_k) + c_l \cdot s_k$. כל השחקנים שמגלים אי התאמה מתלוננים כנגד D. אם יש יותר מ־t שחקנים שהתלוננו מסיקים כי D רע ופוסלים אותו. אם פחות צ־t שחקנים מתלוננים, d נותן לכולם את הערכים שהוא חילק לשחקנים שהתלוננו. בצעד הבדיקה השני, שוב כל p_l בוחר ביט מקרי אבל כעת נעזר ב־r הפולינומים האחרונים ש־D חילק. כולם בודקים את הערכים שלהם אבל גם את הערכים הפומביים. p_k מתלונן כנגד D אם יש אי התאמה. אן בסה"כ היו למעלה מ־t תלונות, D נפסל. אחרת, כל שחקן לוקח את הערכים האחרונים שקיבלת מ־D עבור הפולינום $f(x)$.
נותר להראות: אם מספר המתללוניים בשני השלבים כנגד Dהוא לכל היותר t אזי בהסתברות גבוהה הערכים שבידי השחקנים הטובים של הפולינום $f(x)$נמצאים על פולינום ממעלה $\geq t$. **הוכחה:** ולכן יש $n \geq 3t+1$ $t+1$ שחקנים שלא התלוננו. נסמן קבוצה זו ב־$G \subseteq n$ ($|G| = t+1$). נגדיר פולינומים $f(x), g_1(x), ..., g_{2r}(x)$ ממעלה לכל היותר t שמתאימים לערכים שבידי G. נראה שבהסתברות גבוהה כל השחקנים הטובים מחזיקים ערכים מעל $f^G(x)$. נניח שלא, אז יהיה שחקן p כך ש־$f(\alpha_p) \neq f^G(\alpha_P)$. נניח ש־$b_k$ נבחר ע"י שחקן טוב. נתבונן ב־2 מקרים:

1. $g_k(\alpha_p) = g_k^G(\alpha_p)$.
2. $g_k(\alpha_p) \neq g_k^G(\alpha_p)$.

אם אנו במקרה 2 ו־$b_k = 0$ אזי p מתלונן. D מגלה את $h_k(x) = g_k^G(\alpha_p)$ אחרת מישהו מ־G היה מתלונן. אם אנו במקרה 1 אזי הפולינום $h_k(x) = f^G(x) + g_k^G(x)$ אחרת מישהו מ־G היה מתלונן ו־$b_k = 1$. אם קורה מקרה 2 ו־$b_k = 1$ אנו לא יודעים דבר. $h_k(\alpha_p) \neq g_k(\alpha_p) + f(\alpha_p)$ ולכן p היה מתלונן. לכן, בהסתברות $1/2$, אם מישהו לא היה על הפולינום, הוא היה מתלונן. אם הערכים ש־D מגלה למתלוננים אינם על הפולינומים שנבעו ע"י G בשלב השני בהסתברות גבוהה. אזי, עבור כל ביט מקרי טוב יש התסברות $1/2$ שכל השחקנים ב־G התלוננו ולכן הסתברות השגיאה היא (כשלא תפסנו שחקן רע) $1/2^\ell >$ כש־ℓ שווה למספר הביטים המקרים שנבחרו ע"י שחקנים טוב. (מכיוון שהנחנו ש־G קבוצה טובה אבל יש התנייה שיכול להיות b_k תלוי בבחירת G_k). מכיוון שבחרנו מראש את G, צריך לדאוג לכך שההסתברות השגיאה תהיה קטנה גם כאשר מסכמים את כל הבחירות האפשריות של G. אם $\ell \geq n + log\, \frac{1}{\epsilon}$ אזי הסתברות השגיאה לכל G קטנה מ־$\epsilon/2^n$ ולכן קטנה מ־ϵ. ■

חישוב בטוח: נזכיר כי $t < n/3$. כל שחקן מקבל בחלוקת סוד את הקלטים. פעולות ליניאריות כמו קודם בפרט אם רוצים ביחד לייצר איבר מקרי בשדה: כל שחקן p_k מחלק r_k מקרי בשדה, מחברים את הסכום $r = \Sigma_{i=1}^{n} r_i$. יש שחקנים טובים שחיקלו וזה בת"ל במה שהשחקנים הרעים עשו.

נתאר פרוטוקול חלוקת סוד בין שלשה a, b, c. a, b, c מחלקים בדרך הרגילה (בחלוקת סוד t-בטוחה) ובנוסף המחלק D יוכיח לכולם שמתקיים: $a \cdot b = c$. לצורך הבדיקה D בוחר מספרים מקריים בשדה: $r_{1,...,}r_{2k}, s_{1,...,}s_{2k}$ ומחשב את

$$(*) \quad d_\ell = (a + r_\ell) \cdot (b + s_\ell)$$

D מחלק בתור סודות את r_ℓ, s_ℓ, d_ℓ עבור $\ell = 1, ..., 2k$. השחקנים בוחרים יחד קבוצה מקרית $y = \{j_1, ..., j_k\}$ כאשר $y \subseteq \{1, ..., 2k\}$. עבור האינדקסים $\ell \in y$ נבדוק את $(*)$: נחשב $a + r_\ell, b + s_\ell$ (זה בטוח כי זה עדיין מקרי) ונגלה את d_ℓ לכולם ובודקים. אנו יודעים כי $a \cdot b = c$ צריך להיו נכון. אם $\ell \notin y$ נגלה את r_ℓ, s_ℓ

$$(**) \quad d_\ell = a \cdot s_\ell + b \cdot r_\ell + a \cdot b + r_\ell \cdot s_\ell$$

..אנחו לא יודעים את: O

ניקח ונחסיר $(**) - d_\ell$ ונקבל פולינום התוצאה וננודא שאכן איבר החופשי שלו הוא 0. אם אכן $(*)$ נכון אזי $a \cdot b = c$. האפשרות היחידה לכך ש־$a \cdot b \neq c$ היא קטנה או שווה ל־$\frac{1}{\binom{2k}{k}} < \frac{1}{2^k}$.

חישוב t-בטוח ל־n שחקנים ($t < n/3$): כל שנותר להראות זה פרוטוקול שער כפל:

a חולק בחלוקת סוד (טובה) בעזרת $A(x)$. b חולק בחלוקת סוד טובה בעזרת $B(x)$. שחקן p_k יכול לחשב את $d_k = A(\alpha_k) \cdot B(\alpha_k)$. נבקש מ־$p_k$ לחלק את $A(\alpha_k)$ ואת $B(\alpha_k)$ואת d_k ולהוכיח שאכן $c = j$ מתקיים. החלקים $A(\alpha_1), ..., A(\alpha_n)$ נמצאים על הפולינום ממעלה לכל היותר t. אבל מבין ערכים שקיבלו בפועל יכולים להיות עד t שגיאות. נחשב בעזרת חישוב לינארי את וקטור הסינכרון של השגיאה. זה $2t$ קומבינציות ליניאריות שאמורת להיות אפס אם אין שגיאות ובעזרתן ניתן לחשב את מיקום השגיאות וגודלן.

5 התחייבות על ביט

נתאר את המודל: A, B הן מכונות פולינומיאליות. A מחזיקה קלט $b \in \{0, 1\}$. לפרוטוקול שני שלבים:

בשלב א' ישנה התחייבות של A בפני B. בסוף השלב B מחזיקה אינפורמציה כך שמתקיים:

1. B אינו יכול להבחין (חישובית) בין המקרה $b = 0$ לבין המקרה $b = 1$. אם מגדירים את פרמטר הבטיחות k, כוונתנו היא שלכל אלג' פולינומיאלי ב־k D ולכל פולינום $p(k)$ ההסתברות ש־D, בהינתן הקלט שבידי B, מצליח להבחין בין המקרים קטנה מ־$1/p(k)$ עבור k מספיק גדול. דרך אחרת להגדיר זאת היא על ידי מונחי (s, ϵ) כפי שעשינו בעבר.
2. אם B איננו פוסל את A בשל התנהגות לא ראויה אזי יש ערך יחיד b אותו A (המוגבלת בזמן פולינומיאלי) יכולה לגלות בהמשך שתואם את ההתחייבות שלה.

בשלב ב' A מגלה ל־B את האינפורמציה ל־B כך שיוכל לפתוח ולגלות.

נתאר מספר פרוטוקולים המקיימים את הדרישות הנ"ל.

5.1 פרוטוקול ראשון

תהי $f = \{f_k\} : \{0,1\}^k \to \{0,1\}$ פרמוטציה חד כיוונית עם ביט קשה $h = \{h_k\} : \{0,1\}^k \to \{0,1\}$. A בוחרת $x \in \{0,1\}^k$ מקרי ומחשבת את $f_k(x)$. ושולחת את ההתחייבות $m_b = (f_k(x), h_k(x) \oplus b)$. מהגדרת ביט קשה B לא יכול להבחין בין m_b לבין U_{k+1}. בשלב הגילוי A שולחת ל־B את x ואת b. B בודק שאכן הפרמטרים מתאימים. A לא יכולה לרמות מכיוון ש־f פרמוטציה ולכן יש רק x יחיד מתאים. בפרוטקול זה ההתחייבות של A מוחלטת. חוסר הידיעה של B הינו חישובי.

5.2 פרוטוקול שני

בפרוטוקול הראשון הנחנו קיום פרמוטציה. זאת הנחה חזקה. נראה פרוטוקול אחר שמניח משהו חלש יותר - קיום פונ' חד־כיוונית אשר באמצעותה נוכל לייצר יוצר פ"א כפי שלמדנו בעבר. (פונ' חד כיוונית היא הנחה חלשה יותר במובן הזה שלא ידוע איך בהינתן פונ' חד־כיוונית בונים פרמוטציה חד־כיוונית.)

בפרוטוקול השני אנחנו מניחים קיום יוצר פ"א $G : \{0,1\}^k \to \{0,1\}^{3k}$. B בוחר סדרה מקרית $(r_1, ..., r_{3k}) \in U_{3k}$ ושולח ל־A. A בוחרת $x \in U_k$ ומחשבת

$$m_b = \begin{cases} G(x) & \text{if } b = 0 \\ g(x) \oplus r & \text{if } b = 1 \end{cases}$$

ושולחת ל־B. (אם B היה יכול לבחור r כך שיוכל להבחין בין $G(x)$ ל־$G(x) \oplus r$ אזי היה יכול גם להבחין בין $g(x)$ ל־U_{3k})

בשלב הגילוי A שולחת את x ו־b ו־Bבודק. A יכולה לרמות בהסתברות שגיאה אקספוננציאלית: נשים לב שעל מנת ש־A תוכל לרמות במקרה ש־B שלח את הסדרה r, צריך שיהיו x, y ב־$\{0,1\}^k$ כך ש־$r = g(x) \oplus g(y)$. באופן שקול, צריך שיהיו $x, y \in \{0,1\}^{3k}$ כך ש־$r = g(x) \oplus g(y)$חישובית, כך שאם קיים A כנ"ל היא תוכל לאתר אותו. אנו כן נוכל להגביל אותה סטטיסטית (מנגד, B מוגבל חישובית ולא יכול לבצע זוג x, y בדיקות שכאלה). קיימים (פחות מ־) 2^{2k} זוגות שונים x, y כנ"ל. לכן, התמונה של הפונקציה $f(x, y) = g(x) \oplus g(y)$ מכילה (פחות מ־) 2^{2k} איברים. מנגד, B בוחר איבר r בצורה אחידה מתוך 2^{3k} איברים. לכן, הסיכוי שקיימים $x, y \in \{0,1\}^{3k}$ כך ש־$r = g(x) \oplus g(y)$ קטן מ־$\frac{2^{2k}}{2^{3k}} = \frac{1}{2^k}$ וזוהי הסתברות אקספוננציאלית קטנה ב־k.

5.3 פרוטוקול שלישי

בפרוטוקולים הקודמים הנחנו ש־B מוגבל חישובית ומכאן נובע חוסר הידיעה שלו. מנגד, הנחנו ש־A אינה מוגבלת חישובית. בפרוטוקול הראשון הצלחנו למנוע מ־A את האפשרות לרמות בעזרת שימוש בפרמוטציה. בפרוטוקול השני, A יכולה לרמות בהסתברות שגיאה קטנה אקספוננציאלית. כעת, נתאר פרוטוקול בו B אינו מוגבל חישובית אך לא מקבל שום מידע. מנגד, A מוגבלת חישובית ומכאן מוגבלת ביכולתה לרמות. נציין שאפשר לממש פרוטוקול שכזה באמצעות פונ' חד־כיווניות. אנו נתאר פרוטוקול הנשען על תורת המספרים. המנגנון שנתאר ישען במידה רבה על סכימת ההצפנה של רבין. בדרכנו נזדקק למספר כלים בתורת המספרים.

הערה 5.1 לא כל ההוכחות ניתנו בכיתה. ניתן לקבל כל אחת מהטענות בתורת המספרים כפי שהיא ולהתקדם לתיאור הפרוטוקולים. השתדלתי שכן להביא את ההוכחות המלאות (נעזרתי בספר של $katz\&Lindell$).

5.3.1 $RSA - reminder$

נזכיר את סכימת RSA:

1. בוחרים p, q ראשוניים מקריים ושונים. נסמן $N = p \cdot q$.
2. בוחרים e כך ש־$gcd(e, \phi(n)) = 1$. נזכיר כי $\phi(n) = (p-1) \cdot (q-1)$.
3. פותרים $d = e^{-1}(mod\, \phi(N))$.
4. המפתח הציבורי הינו e והמפתח הפרטי הינו d . נשים לב כי עבור הודעה m מתקיים:

$$(m^e)^d = m^{k \cdot \phi(N)+1} = m^{(p-1)\cdot(q-1)+1} = m \cdot (m^{p-1})^{q-1} = m \cdot 1^{q-1} = m$$

כנדרש.

סכימה זו מסתמכת על האמונה הרווחת לפיה קשה לפרק מספר לגורמיו הראשוניים (למעשה, הסכימה מסתמכת על האמונה הרווחת לפיה קשה לחשב את $\phi(N)$ כאשר הפירוק של N אינו ידוע. עם זאת, הבעיות הנ"ל שקולות. כלומר, פירוק N ומציאת $\phi(N)$ קשים באותה מידה כאשר הפירוק של N אינו ידוע). חשוב לציין, כי לא הראו עד היום כי פיענוח RSA ופירוק N קשים באותה מידה. כמובן, שאם RSA קשה אזי גם פירוק N קשה. עם זאת, לא קיימת הוכחה שלא ניתן לפתור את בעית ה־RSA מבלי לפרק N (או לחילופין למצוא את $\phi(N)$ או d). אנו נתאר בהמשך את סכימת ההצפנה של רבין, עבורה קיימת הוכחה שכזו. לשם כך נזדקק למספר כלים בתורת המספרים.

5.3.2 מציאת שורשים ריבועיים ב־$\mathbb{Z}_p^*$

הערה 5.2 נזכיר את **משפט פרמה הקטן:** לכל p ראשוני, ולכל מספר שלם a מתקיים $a^{p-1} = 1$.

הערה 5.3 נזכיר את משפט השאריות הסיני: יהיו p, q שני מספרים ראשוניים (למעשה מספיק ש־p, q זרים זה לזה). נסמן $N = p \cdot q$. אזי,

$$\mathbb{Z}_N \backsimeq \mathbb{Z}_p \times \mathbb{Z}_q \qquad and \qquad \mathbb{Z}_N^* \backsimeq \mathbb{Z}_p^* \times \mathbb{Z}_q^*$$

יתרה מכך, ההעתקה f הממפה איברים $x \in \{0, ..., N-1\}$ לזוגות (x_p, x_q) עבור $x_p \in \{0, ..., p-1\}, x_q \in \{0, ..., q-1\}$ המוגדרת כך:

$$f(x) = (x \,(mod\, p), x \,(mod\, q))$$

הינה איזומורפיזם מ־$\mathbb{Z}_N$ ל־$\mathbb{Z}_p \times \mathbb{Z}_q$ וכמו כן איזומורפיזם מ־$\mathbb{Z}_N^*$ ל־$\mathbb{Z}_q^* \times \mathbb{Z}_p^*$. משמע,

1. f הינה חח"ע (וגם 'על' במקרה זה בו החבורות המדוברות הינן סופיות).
2. לכל $x, x' \in \mathbb{Z}_N$ מתקיים $f(x + x') = f(x) + f(x')$.
3. לכל $x, x' \in \mathbb{Z}_N^*$ מתקיים $f(x \cdot x') = f(x) \cdot f(x')$.

הערה 5.4 במספר מקרים לאורך הדרך אנו נניח כי המספרים הראשוניים p, q בהם נשתמש מקיימים $p \equiv q \equiv 3 \,(mod\, 4)$. הנחה זו סבירה בהתחשב במשפט הצפיפות של דריכלה. מה עוד, שלצרכי סכימת ההצפנה של רבין, זה יספיק לנו.

א אנו נראה כעת כיצד, בהנתן p ראשוני, ניתן לקבוע לכל $a \in \mathbb{Z}_p^*$ האם הוא ריבוע ב־$\mathbb{Z}_p^*$. לאח"כ נראה כיצד ניתן למצוא את שורשיו במידה ומצאנו כי a אכן ריבוע.

הגדרה 5.5 נסמן:

$$\mathcal{QR}_p = \{y \in \mathbb{Z}_p^* : \exists x \in \mathbb{Z}_p^* \, s.t \quad x^2 = y \,(mod\, p)$$

$$\mathcal{QNR}_p = \{a \in \mathbb{Z}_p^* : a \notin \mathcal{QR}_p\}$$

במילים, $\mathcal{QR}_p$ היא קבוצת הריבועים ב־$\mathbb{Z}_p^*$ ו־$\mathcal{QNR}_p$ היא משלימתה.

טענה 5.6 יהי $p > 2$ מספר ראשוני. לכל ריבוע בשדה קיימים בדיוק שני שורשים.

הוכחה: יהי $y \in \mathbb{Z}_p^*$ לגביו אנו יודעים כי y מהווה ריבוע. נסמן ב־$x \in \mathbb{Z}_p^*$ את האיבר המקיים $y = x^2 \,(mod\, p)$. כלומר, x הינו שורש. ברור כי גם $-x$ הוא שורש, שכן $(-x)^2 = x^2 = y \,(mod\, p)$. בנוסף, ברור כי $-x \neq x \,(mod\, p)$. שכן, אחרת היה מתקיים $x + x = 0 \,(mod\, p)$ ואז היינו מסיקים כי p זוגי או ש־p מחלק את x בסתירה להנחותינו. אם כן, הראינו כי קיימים לפחות 2 שורשים שונים. נראה שלא יתכנו שורשים נוספים. נניח בשלילה כי קיים $x' \neq \pm x \,(mod\, p)$ כך ש־$x'^2 = y \,(mod\, p)$.אזי, מתקיים $x^2 - x'^2 = 0 \,(mod\, p)$. באופן שקול, מתקיים $(x - x') \cdot (x + x') = p \,(mod\, p)$. מכאן, בהכרח $p|(x - x')$ או $p|(x + x')$. במקרה הראשון, נסיק $(x' = x) \,mod\, p$ בסתירה. במקרה השני, נסיק $x' = -x \,(mod\, p)$ בסתירה. ■

מסקנה 5.7 עבור $p > 2$ ראשוני, מתקיים $|\mathcal{QR}_p| = |\mathcal{QNR}_p| = |\mathbb{Z}_p^*|/2 = (p-1)/2$.

הוכחה: נובע מיידית מהטענה הקודמת, שכן, הראינו כי ההעתקה $f : \mathbb{Z}_p^* \to \mathbb{Z}_p^*$ המוגדרת ע"י $f(x) = x^2 \,(mod\, p)$ היא $2 \to 1$. ■

כעת, נאפיין מיהם הריבועים ב־$\mathbb{Z}_p^*$.

הגדרה 5.8 **Jacobi Symbol** עבור $p>2$ ראשוני, נגדיר לכל $x \in \mathbb{Z}_p^*$:

$$\mathcal{J}_p = \begin{cases} +1 & x \text{ is a quadratic residue modulo } p \\ -1 & x \text{ is not a quadratic residue modulo } p \end{cases}$$

הגדרה זו ניתנת להרחבה לכל xזר ל־p בצורה הבאה:

$$\mathcal{J}_p(x) - \mathcal{J}_p(x \bmod p)$$

הערה 5.9 למעשה זהו $Legendre\,Symbol$ וההכללה שנראה בהמשך ל־N עבור $N = p \cdot q$ נקראת $Jacobi\,Symbol$. על מנת להצמד למינוח בספר הלימוד ($Katz\&Lindell$) אנו נשתמש במינוח סימבול יעקובי בשני המקים ונבין מן ההקשר באיזה מהם מדובר.

טענה 5.10 יהי $p>2$ מספר ראשוני. אזי, $\mathcal{J}_p(x) = x^{\frac{x-1}{2}}$.

הוכחה: נזכר כי $\mathbb{Z}_p^*$ היא חבורה ציקלית. אם כך, יהי g יוצר של החבורה.
נניח תחילה כי x הוא ריבוע. נרצה להראות כי $x^{\frac{p-1}{2}} = 1$. מכיוון ש־x הוא ריבוע קיים i זוגי כך ש־$x = g^i$. נסמן $i = 2j+1$. מתקיים:

$$x^{\frac{p-1}{2}} = (g^{2j})^{\frac{p-1}{2}} = (g^{p-1})^j = 1^j$$

כאשר המעבר הלפני אחרון נובע מהמשפט הקטן של פרמה.
כעת, נניח כי x אינו ריבוע. אזי, קיים i אי זוגי כך ש־$g^i = x$. נסמן $i = 2j+1$. מתקיים:

$$x^{\frac{p-1}{2}} = (g^{p-1})^j \cdot g^{\frac{p-1}{2}} = g^{\frac{p-1}{2}}$$

כעת, נשים לב כי $(g^{\frac{p-1}{2}})^2 = g^{p-1} = 1$ לפי המשפט הקטן של פרמה. לכן, $g^{\frac{p-1}{2}} = \pm 1$. נשתכנע כי $g^{\frac{p-1}{2}} \neq 1$ ומכאן ינבע כי $g^{\frac{p-1}{2}} = -1$. ניזכר כי g יוצר.g אם כך הוא מסדר $p-1$. מכאן, לא יתכן כי $g^{\frac{p-1}{2}} = 1$. ■

מסקנה 5.11 בהנתן מספר ראשוני $p>2$, ניתן לקבוע בזמן פולינומיאלי לכל $x \in \mathbb{Z}_p^*$ אם x הינו ריבוע.

הוכחה: נובע ישירות מהעובדה כי הראינו דרך לחישוב סימבול יעקובי של x ע"י העלאה בחזקה (העלאה בחזקה מודולו p ניתנת למימוש בזמן פולינומיאלי). ■

כעת, נראה דרך לחישוב השורשים של ריבוע ב־$\mathbb{Z}_p^*$.
נתאר מספר תכונות של סימבול יעקובי שנובועות די בקלות מהגדרות ומהתכונות שכבר הראינו.

טענה 5.12 יהי $p>2$ ראשוני. לכל $x, y \in \mathbb{Z}_p^*$ מתקיים

$$\mathcal{J}_p(x \cdot y) = \mathcal{J}_p(x) \cdot \mathcal{J}_p(y)$$

מסקנה 5.13 יהי $p>2$ ראשוני. יהיו $x, x' \in \mathcal{QR}_p$ ו־$y, y' \in \mathcal{QNR}_p$. מתקיים:
1. $x \cdot x' \,(mod\, p) \in \mathcal{QR}_p$
2. $y \cdot y' \,(mod\, p) \in \mathcal{QR}_p$
3. $x \cdot y \,(mod\, p) \in \mathcal{QNR}_p$

טענה 5.14 יהי $p > 2$ מספר ראשוני וכמו כן אנו נניח כי $p \equiv 3 \,(mod\, 4)$. נתון $a \in \mathbb{Z}_p^*$ ריבוע. אזי, $a^{\frac{p+1}{4}}$ הינו שורש של a (והשורש השני של a הינו $-(a^{\frac{p+1}{4}})$).

הוכחה: תחילה, נשים לב כי מהנתחותינו נובע כי $(p+1)/4$ הוא מספר שלם. נסמן $p = 4i + 3$. מתקיים:

$$(a^{\frac{P+1}{4}})^2 = a^{\frac{p+1}{2}} = a^{\frac{p-1}{2}} \cdot a = 1 \cdot a = a$$

כאשר המעבר הלפני אחרון נובע מכך ש־a הינו ריבוע ומהגדרת סימבול יעקובי ותכונותיו. ■

הערה 5.15 קיים גם אלגוריתם (הסתברותי) למציאת השורשים של a בחבורה $\mathbb{Z}_p^*$ במקרה $p \equiv 1 \,(mod\, 4)$. אנו לא נתאר אותו וגם לא נזדקק לו בהמשך.

ניעזר בידע שלנו למצוא שורשים ב־$\mathbb{Z}_p^*$ עבור p ראשוני על מנת למצוא שורשים ב־$\mathbb{Z}_N^*$.

5.3.3 מציאת שורשים ריבועיים ב־$\mathbb{Z}_N$ עבור $N = p \cdot q$ כאשר p, q ראשוניים והפירוק ידוע

בחלק זה נשלב את ההתוצאות אליהן הגענו בחלק הקודם עם משפט השאריות הסיני על מנת לאפיין את השורשים ב־$\mathbb{Z}_N^*$. נסמן ע"י $\leftrightarrow$התאמה בין איבר ב־$\mathbb{Z}_N^*$ למקבילו ב־$\mathbb{Z}_q^* \times \mathbb{Z}_p^*$. המשפט הבא יעזור לנו לאפיין מיהם הריבועים ב־$\mathbb{Z}_N^*$.

טענה 5.16 יהי $N = p \cdot q$ עבור p, q ראשוניים, אי זוגיים שונים זה מזה. יהי $y \in \mathbb{Z}_N^*$ ונניח כי $y \leftrightarrow (y_p, y_q)$. אזי, y הינו ריבוע ב־$\mathbb{Z}_N^*$ אם"ם y_p הינו ריבוע ב־$\mathbb{Z}_p^*$ ו־y_q הינו ריבוע ב־$\mathbb{Z}_q^*$.

הוכחה: נניח כי y הוא ריבוע מודולו N. נסמן ב־x את האיבר בחבורה המקיים $x^2 = y \,(mod\, N)$. יהיו $x_p \in \mathbb{Z}_p^*, x_q \in \mathbb{Z}_q^*$ כך שמתקיים $x \leftrightarrow (x_p, x_q)$.
לפי מפשט השאריות הסיני מתקיים:

$$(y_p, y_q) \leftrightarrow y = x^2 \leftrightarrow (x_p, x_q)^2 = (x_p^2 \,(mod\, p), x_q^2 \,(mod\, q))$$

ומכאן $y_p = x_p^2 \,(mod\, p), y_q = x_p^2 \,(mod\, q)$.
את הכיוון השני מראים בצורה דומה. ■

בירור מעמיק יותר מראה כיצד ניתן לאתר את השורשים של ריבוע ב־$\mathbb{Z}_N^*$.

טענה 5.17 יהי $y \in \mathbb{Z}_N^*$. נניח כי $y \leftrightarrow (y_p, y_q)$ ונניח כי $y_p = x_p^2 \,(mod\, p), y_q = x_p^2 \,(mod\, q)$.אזי, ל־$y$ 4 שורשים ב־$\mathbb{Z}_N^*$ ואלו הם האיברים ב־$\mathbb{Z}_N^*$ המתייחסים ל:

$$(\pm x_p, \pm x_q)$$

הוכחה: ראשית, לפי טענה נשים לב כי לפי טענה 5.16 y הינו ריבוע ב־N. כעת, מתקיים:

$$(\pm x_p, \pm x_q)^2 = (x_p^2 \,(mod\, p), x_q^2 \,(mod\, q)) = (y_p, y_q) \leftrightarrow y$$

לכן, 4 האיברים המתייחסים ל־$(\pm x_p, \pm x_q)$ הם שורשים ריבועיים של y ב־$\mathbb{Z}_N^*$. נציין כי משפט השאריות הסיני אכן מבטיח שמדובר ב־4 איברים שונים. ■

הערה 5.18 נשתמש בסימונים $\mathcal{QR}_N, \mathcal{QNR}_N$ בצורה סימטרית לצורה בה השתמשנו בסימונים אלו עבור p, q.

טענה 5.19 ייהי $N = p \cdot q$ עבור p, q ראשוניים, אי זוגיים שונים זה מזה. מתקיים:

$$|\mathcal{QR}_N| = |\mathbb{Z}_N^*/4| = \frac{(p-1)\cdot(q-1)}{4}$$

הוכחה: נובעת ישירות מטענה 5.16 שכן, הראינו כי ההעתקה $f : \mathbb{Z}_N^* \to \mathbb{Z}_N^*$ המוגדרת ע"י $f(x) = x^2 \,(mod\, p)$ היא $4 \to 1$.

כעת, נכליל את סימבול יעקובי. ■

הגדרה 5.20 יהי $N = p \cdot q$ עבור p, q ראשוניים שונים זה מזה. עבור x זר ל־N נגדיר:

$$\mathcal{J}_N(x) = \mathcal{J}_p(x) \cdot \mathcal{J}_q(x)$$

הגדרה 5.21 ייהי $N = p \cdot q$ עבור p, q ראשוניים, אי זוגיים שונים זה מזה. נגדיר ע"י $\mathcal{J}_N^{+1}$ את קבוצת האיברים ב־$\mathcal{J}_N$ עבורם סימבול יעקובי שווה ל־1. מנגד, נגדיר ע"י $\mathcal{J}_N^{-1}$ את קבוצת האיברים ב־$\mathcal{J}_N$ עבורם סימבול יעקובי שווה ל־-1.

במקרה של $\mathbb{Z}_p^*$ היתה לנו את התכונה הבאה: x הינו ריבוע אם"ם $\mathcal{J}_p(x) = 1$ (למעשה זאת היתה ההגדרה). נשים לב שתכונה זו לא נשמרת במלואה ב־$\mathbb{Z}_N^*$. כיוון אחד נשמר; אם x הינו ריבוע אזי לפי טענה 5.16 ולפי הגדרת סימבול יעקובי המקורית מתקיים $\mathcal{J}_p(x) = \mathcal{J}_q(x) = 1$. כעת, לפי ההכללה של סימבול יעקובי,מתקיים $J_N(x) = 1$. מנגד, נחשוב על x כך ש־$J_q(x) = \mathcal{J}_p(x) = -1$. במקרה זה מתקיים $\mathcal{J}_N(x) = 1$ על אף כי ברור ע"פ טענה 5.16 כי x אינו ריבוע. לשם כך נגדיר:

הגדרה 5.22 יהי $N = p \cdot q$ עבור p, q ראשוניים, אי זוגיים שונים זה מזה. נגדיר ע"י $\mathcal{QNR}_N^{+1}$ את אותם איברים ב־$\mathbb{Z}_N^*$ אשר להם סימבול יעקובי 1 אך אינם ריבועים.

מבט מעמיק בתוצאות אליהן הגענו עד כה מספק בצורה די ישירה את המסקנות הבאות:

מסקנה 5.23 יהי $N = p \cdot q$ עבור p, q ראשוניים שונים זה מזה. מתקיים:
1. $\mathcal{J}_N^{+1} = |\mathbb{Z}_N^*|/2 = \frac{(p-1)\cdot(q-1)}{2}$
2. $\mathcal{QR}_N \subseteq \mathcal{J}_N^{+1}$.
3. $|\mathcal{QR}_N| = |\mathcal{J}_N^{+1}|/2$

תכונה נוספת הקלה להוכחה וסימטרית לתכונה אותה הראינו ב־$\mathbb{Z}_p^*$ היא התכונה הבאה:

טענה 5.24 יהי $N = p \cdot q$ עבור p, q ראשוניים, אי זוגיים שונים זה מזה. לכל $x, y \in \mathbb{Z}_N^*$ מתקיים:

$$\mathcal{J}_N(x \cdot y) = \mathcal{J}_N(x) \cdot \mathcal{J}_N(y)$$

בצורה די ישירה מתקבלות המסקנות הבאות:

מסקנה 5.25 יהי $N = p \cdot q$ עבור p, q ראשוניים, אי זוגיים שונים זה מזה. יהיו $x, x' \in \mathcal{QR}_N$ ו־$y, y' \in \mathcal{QNR}_N^{+1}$. מתקיים:
1. $x \cdot x' \,(mod\, N) \in \mathcal{QR}_N$
2. $y \cdot y' \,(mod\, N) \in \mathcal{QR}_N$
3. $x \cdot y \,(mod\, N) \in \mathcal{QNR}_N^{+1}$

נשים לב שבניגוד למסקנה 5.13 , כאן לא מובטח שאם $x, x' \in \mathcal{QNR}_N$ אזי $x \cdot x' \in \mathcal{QR}_N$. למשל אם מתקיים $\mathcal{J}_p(x) = \mathcal{J}_q(x') = 1$ ו־$\mathcal{J}_q(x) = \mathcal{J}_p(x') = -1$ אזי ברור כי $x, x' \in \mathcal{QNR}_N$ אך נשים לב כי גם $x \cdot x' \in \mathcal{QNR}_N$.
לאחר שסיימנו לאפיין מיהם הריבועים ב־$\mathbb{Z}_N^*$ נציע שיטה לאתרם (כאשר הפירוק של N ידוע).

נניח כי נתון $a \in \mathbb{Z}_N^*$, כך ש־$a \leftrightarrow (a_p, a_q)$. הראינו בחלק הקודם כיצד ניתן בזמן פולינומיאלי לבדוק אם a_p הינו ריבוע ב־$\mathbb{Z}_p^*$ ו־a_q הינו ריבוע ב־$\mathbb{Z}_q^*$. נניח כי a_p ריבוע וכך גם a_q. אז לפי טענה 5.16 , a הוא ריבוע ב־$\mathbb{Z}_N^*$. יתרה מכך, לפי טענה 5.17 השורשים של a הם 4 האיברים ב־$\mathbb{Z}_N^*$ המתאימים ל־$(\pm a_p, \pm a_q)$. נשאלת השאלה כיצד ניתן למצוא ביעילות את אותם 4 איברים. נתבונן בתכונות האיזומורפיזם $f : \mathbb{Z}_N^* \to \mathbb{Z}_p^* \times \mathbb{Z}_q^*$ שכזכור מוגדר ע"י $f(x) = (x \,(mod\, p), x \,(mod\, q))$. ברור כי בהנתן $x \in \mathbb{Z}_N^*$ קל לחשב את $f(x)$. מסתבר שקיימת גם דרך יעילה לחשב את $f^{-1}(x)$. כלומר, בהנתן (x_p, x_q) נוכל למצוא את x בצורה יעילה. (לא נפרט כאן אודות הנושא. למי שמתעניין קיים פירוט ב־$Katz\&Lindell\ section\ 7.1.5$).

5.3.4 מציאת שורשים ריבועיים ב־$\mathbb{Z}_N$ עבור $N = p \cdot q$ כאשר p, q ראשוניים והפירוק אינו ידוע

בחלק הקודם הראינו כיצד ניתן למצוא שורשים ריבועיים מודולו N בצורה יעילה כאשר הפירוק ידוע. אנו נראה כעת, שכאשר הפירוק אינו ידוע, מציאת שורשים מודולו N קשה בדיוק כמו פירוק N. כיוון אחד הוא טריוויאלי למדי; אם ניתן לפרק ביעילות את N אזי לפי התוצאות מהחלק הקודם, ניתן למצוא ביעילות שורשים מודולו N. מכאן, פירוק N קשה לפחות כמו מציאת שורשים. נראה כעת את הכיוון השני.

טענה 5.26 יהי $N = p \cdot q$ עבור p, q ראשוניים, א"ז ושונים זה מזה. בהנתן x, x' כך ש־$x \neq \pm x' \,(mod\, N)$ ו־$x^2 = y = x'^2 \,(mod\, N)$ ניתן לפרק את N בזמן פולינומיאלי.

הוכחה: אנו טוענים כי אחד מבין $gcd(N, x + x')$ו־$gcd(N, x - x')$ שווים לאחד מבין p ו־q. ברור כי אם טענה זו נכונה ניתן לבדוק את שתי האפשרויות וכך למצוא p, q בזמן פולינומיאלי (gcd ניתן לחישוב בזמן פולינומיאלי). מכיוון ש־$x^2 = x'^2 \,(mod\, N)$ מתקיים:

$$0 = x^2 - x'^2 = (x + x') \cdot (x - x') \,(mod\, N)$$

ומכאן $N|(x + x') \cdot (x - x')$. אזי $p|(x + x') \cdot (x - x')$. מכאן $p|(x + x')$ או $p|(x - x')$. נניח כי $p|(x + x')$ (ההוכחה דומה במקרה הנגדי). אם $q|(x + x')$ אזי $N|(x + x')$אבל במצב זה $x = -x' \,(mod\, N)$ בסתירה להנחה. מכאן $q \nmid (x + x')$ ולכן $gcd(N, x + x') = p$. ■

מהטענה האחרונה מתקבל בקלות המשפט הבא:

משפט 5.27 חישוב שורשים ריבועיים מודולו N קשה לפחות כמו פירוק N.

הוכחה: ההוכחה כאמור נובעת ישירות מן הטענה. בספר קיים תיאור של אלגוריתם לפירוק N שמשתמש באלגוריתם נתון לחישוב שורשים כמעין "קופסא שחורה" (פרק 11.2.2 ב־$Katz\&Lindell$). ■

המשפט האחרון מספק לנו משפחה של פונ' חד כיווניות (מתוך הנחה שפירוק N קשה). הפונקציה f_N המתאימה ל־N מקבלת כקלט $x \in \mathbb{Z}_N^*$ (ונניח כי xנדגם בהתפלגות אחידה על פני איברי $\mathbb{Z}_N^*$)ומעתיקה אותו ל־$x^2 \,(mod\, N)$. נראה מקרה מיוחד בו מתקבלת משפחה של פרמוטציות חד כיווניות.

הגדרה 5.28 N הוא $Blum\ integer$ אם p, q אשר מקיימים $N = p \cdot q$ הם שני מספרים ראשוניים ושונים המקיימים $p \equiv q \equiv 3 \,(mod\, 4)$.

כעת המתפח לבניית הפרמוטציה ניתן ע"י הטענה הבאה:

טענה 5.29 יהי N ויהיו p, q כך ש־N הינו $Blum\ integer$ ביחס ל־p, q. אזי לכל ריבוע מודולו N קיים בדיוק שורש אחד המהווה גם הוא ריבוע מודולו N.

הוכחה: תחילה נשים לב שתחת ההנחות הנ"ל -1 אינו ריבוע מודולו p וגם אינו ריבוע מודולו q. נראה עבור p. נניח כי $p = 4i + 3$. אזי:

$$(-1)^{\frac{p-1}{2}} = (-1)^{2i+1} = -1 \,(mod\, p)$$

כלומר $\mathcal{J}_p(-1) = -1$ ומכאן -1 אינו שורש מודולו p.
כעת יהי $y \in \mathbb{Z}_N^*$ כך ש־$y \leftrightarrow (y_p, y_q)$ ו־y הינו ריבוע מודולו Nויהיו שורשיו האלמנטים ב־$\mathbb{Z}_N^*$ המתאימים ל־

$$(x_p, x_q), (x_p, -x_q), (-x_p, x_q), (-x_p, -x_q)$$

אנו טוענים שרק אחד מבין האלמנטים הללו הוא ריבוע מודולו N. נניח למשל כי $\mathcal{J}_p(x_p) = 1, \mathcal{J}_q(x_q) = -1$. אזי לפי טענה 5.12 מתקיים:

$$\mathcal{J}_q(-x_q) = \mathcal{J}_q(x_q) \cdot \mathcal{J}_q(-1) = -1 \cdot -1 = 1$$

ולכן, לפי טענה 5.16 האיבר המתאים ל־$(x_p, -x_q)$ הינו שורש מודולו N. באופן דומה $\mathcal{J}_q(-x_p) = -1$ ולכן, שוב לפי טענה 5.16 אף אחד מהשורשים האחרים אינו מהווה ריבוע מודולו N. ■

כעת, יהי N ויהיו p, q כך ש־N הינו $Blum\, integer$ ותהי ההעתקה $f_N : \mathcal{QR}_N \to \mathcal{QR}_N$ הניתנת ע"י $f_N(x) = x^2 \,(mod\, N)$. נניח גם כי x נדגם בהתפלגות אחידה על פני איברי $\mathcal{QR}_N$ (על מנת לדגום בצורה אחידה מ־$\mathcal{QR}_N$ די שנדגום איבר מ־$\mathbb{Z}_N^*$ ונעלה אותו בריבוע). קל לראות ש־f_N הינה פרמוטציה חד־כיוונית (שוב, תחת ההנחה שקשה לפרק את N).

הערה 5.30 קיים אלגוריתם פולינומיאלי לחישוב $\mathcal{J}_N(x)$ גם כשפירוק אינו ידוע.

5.3.5 סכימת BC הנשענת על הקושי לחשב שורשים מודולו N כאשר הפירוק אינו ידוע

נתאר את הפרוטוקול: A מחזיקה ביט $b \in \{0,1\}$ עליו היא מתחייבת. B בוחר p, q שני מספרים ראשוניים ושונים המקיימים $p \equiv q \equiv 3 \,(mod\, 4)$, מחשב את $N = p \cdot q$ ושולח ל־A. A בוחרת x מקרי מודולו N המקיים $\mathcal{J}_N(x) = -1^b$ ושולח את $y = x^2 \,(mod\, N)$ ל־B. נשים לב ש־y מפולג באופן אחיד על פני $\mathcal{QR}_N$. ברור כי ל־Bאין שום אינפורמציה על b. בשלב הגילוי A שולחת את x, b ל־B. לפי הערה 5.30 B יכול לוודא שאכן $\mathcal{J}_N(x) = -1^b$ וכמובן יכול לוודא שאכן $x^2 = y \,(mod\, N)$. נשים לב כי מאחר ו־$\mathcal{J}_p(x) = \mathcal{J}_q(x) = -1$ (הראינו קודם כי -1אינו שורש מודולו p, q עבור p, q ראשוניים המקיימים $p \equiv q \equiv 3 \,(mod\, 4)$) , מתקיים $\mathcal{J}_N(-x) = \mathcal{J}_N(x) \cdot \mathcal{J}_N(-1) = \mathcal{J}_N(x) \cdot 1 = \mathcal{J}_N(x)$. לכן, היכולת של Aלרמות חסומה ע"י יכולתה למצוא $x' \neq \pm x$ כך ש־$x'^2 = y$. אך לפי הנחתנו A המוגבלת חישובית תצליח לעשות כן בהסתברות זניחה.

6 העברה חסרת אבחנה - Oblivious Transfer (סוכם ע"י שיר פלד)

הגדרה 6.1 המקורית על פי רבין: A מחזיקה קלט $b \in \{0,1\}$ ואז A שולחת את הביט, בהסתברות $\frac{1}{2}$ מתקבל b ובהסתברות $\frac{1}{2}$ מתקבל $\#$ (כלומר כלום) והוא מודע לזה שהוא לא קיבל את b. עם זאת - A אינה יודעת ש b לא עבר כהלכה.

.

הגדרה 6.2 (אלטרנטיבית, נקראת גם Oblivious Delivery) A שולחת $a_0, a_1 \in \{0,1\}$.
ל B יש קלט $b \in \{0,1\}$ - כתוצאה מהפרוטוקול B מקבל את a_b ו A לא יודעת איזה מהם התקבל.
באופן פורמלי זוהי פונקציה שנותנת וקטור תוצאה ל A ול B בהתאמה: $F([a_0, a_1], b) = (\perp, a_b)$ (כאן הכוונה ש A אינה מקבלת כל פלט)

איך מהגדרה אחת עוברים לאחרת? נניח שיש לנו פונקציה OD - איך מייצרים OT?

נגדיר ש A מכניס למכונת ה OD את b ואת 0, כעת B מגריל בהסתברות חצי - ביט, ושולח אותו ל A. כעת בוב יודע שאם הוא שלח 0 אז מה שהוא קיבל זה b ואחרת - הוא קיבל 0 שלא נותן לו שום מידע...

בכיוון השני נניח שיש לנו פונקציה OT ורוצים לקבל OD:

A בוחרת $3k$ ביטים מקריים: $a_1, ..., a_{3k}$ ומעבירה אותם דרך ערוץ OT ל B. מתקבלת מחרוזת של ביטים שחלקם $\#$.

B בוחר שתי קבוצות אינדקסים זרות $I_0, I_1 \subseteq \{1, ..., 3k\}$ באופן מקרי כך ש $|I_0| = |I_1| = k$.

אם $b = 0$ אז B דואג כך שבכל הקואורדינטות ב I_0 לא מופיע $\#$. למעט הסתברות שגיאה אקספונ' קטנה ב k מספר המחיקות d בהעברה מקיים $\left|d - \frac{3}{2}k\right| < \frac{1}{4}k$ ולכן אם ב I_0 אין מחיקות אז ב I_1 יש מחיקות. (באופן שקול אם $b = 1$ נבחר את הקבוצות הפוך ונקבל שב I_1 אין מחיקות וב I_0 יש).

B שולח את I_0 ואת I_1 ל A ואז A מחשבת את $r_0 = A_0 \oplus (\oplus_{i \in I_0} a_i)$ (כאשר שני הביטים ש A מתחיל איתם הם A_0 וA_1) ואת $r_1 = A_1 \oplus (\oplus_{i \in I_1} a_i)$.

6.1 פרוטוקול ל .O.T

בהנחה שקיימת פרמוטצית מלכודת (Trapdoor Permutation) כלומר פרמוטציה חד כיוונית, שבעזרת מפתח אפשר להפוך אותה באלגוריתם יעיל (מפתח אחד המתאים לכל הפונקציה, לא צריך מפתח נפרד לכל קלט).

הכוונה כאן היא שיש לנו אלגוריתם יצירת מפתחות שיוצר (E, D) כך ש E באה מתוך אוסף של פרמוטציות חד כיווניות.

נבהיר זאת: באופן כללי אם יש מפתח לפרמוטציה חד כיוונית, במובן זה שניתן באמצעות המפתח להפוך את הפרמוטציה ביעילות - אזי הפרמוטציה אינה חד כיוונית! אנו נדרוש לכן שיהיה אוסף של פרמוטציות כאילו־חד־כיווניות עם מפתחותיהן, כזה שאם נגריל ממנו אחת מקרית - יהיה קשה מאד רק על סמך הפרמוטציה עצמה - להפוך אותה. כלומר לכל דבר ועניין, אם לא ידוע D, אזי היא חד כיוונית.

ניתן לחשוב על זה גם כך - פירוק לגורמים הוא קשה, אבל בהנתן חלק מהגורמים - זה עשוי להיות קל, כלומר במובן זה הם מפתח.

אם מגרילים מספר אקראי - קשה לפרק אותו לגורמים למרות שיש מפתח שמקל על העניין (למשל, רשימת חלק מהגורמים).

כך גם אנחנו מגרילים פונקציה E שקשה להפוך, למרות שיש D שמקל.

נחזור לאלגוריתם:

A בוחרת זוג מפתחות (E, D) ושולחת את E ל B.

נניח בה"כ

$$D, E : \{0,1\}^k \to \{0,1\}^k$$

ו

$$B : \{0,1\}^k \to \{0,1\}$$

היא ביט קשה עבור הפונקציה E.

נניח שהקלט של בוב הוא $b = 0$, והוא בוחר $r_0, r_1 \in \{0,1\}^k$ באופן מקרי ומגדיר

$$s_0 = E(r_0)$$
$$s_1 = r_1$$

ואם $b = 1$ אז:

$$s_0 = r_0$$
$$s_1 = E(r_1)$$

ושולח את (s_0, s_1) לאליס.

אליס מגדירה:

$$w_0 = a_0 \oplus \underbrace{B\left(D\left(s_0\right)\right)}_{B(r_0)}$$
$$w_1 = a_1 \oplus B\left(D\left(s_1\right)\right)$$

ושולחת את (w_0, w_1) לבוב.
אם $b = 0$ אז B יודע את r_0 ולכן יכול לחשב את $B\left(r_0\right)$ ואז $a_0 = w_0 \oplus B\left(r_0\right)$ והוא יכול לחלץ את a_0 בקלות.
באותו מצב (כלומר עדיין $b = 0$), אם r_1 הוא מקרי אזי $B\left(D\left(s_1\right)\right)$ הוא בעצם $B\left(D\left(r_1\right)\right)$, וזה נראה כמו ביט מקרי לבוב גם בהנתן r_1, כי זו בעצם ההגדרה של ביט קשה.
אם $b = 1$ באופן סימטרי הוא יכול לפענח את a_1 בקלות אך לא את a_0.
אם בוב רוצה לרמות - מספיק שהוא נותן את $E\left(r_i\right)$ בשני המקרים - ואז הוא מקבל דברים שהוא יכול לפענח כרצונו...
אליס בכל מקרה לא מקבלת שום אינפורמציה מכיוון ש s_0, s_1 מפולגים אחיד (ה r_i מוגרלים ו E פרמוטציה).

ההנחה שלנו היא שהשחקנים אמינים (אבל סקרנים) נתונה פונקציה $f : X \times Y \to Z$ המוגדרת בעזרת מעגל בוליאני ובעזרת השערים AND שיסומן $\odot$ ושער XORשיסומן כרגיל ע"י $\oplus$.
כמו במקרה של חישוב בין n שחקנים - נקודד כל קלט ושער במעגל באופן הבא, אם מה שאנחנו רוצים לקודד הוא V:

$$V = V_1 \oplus V_2$$

כך ש V_1 מקרי והשני הוא ה XOR שלו, וניתן V_1 לשחקן אחד ו V_2 לשחקן השני.
מבחינת השחקנים - הביטים שבידיהם מוגרלים מקרית.
כאשר A מחזיקה כקלט את x ו B מחזיק כקלט את y, נרצה לחשב את $f\left(x, y\right)$ כך שהפלט יהיה ידוע רק ל B.

שלב הקלט:

לכל ביט a של קלט של A - A בוחרת a_1מקרי ומגדירה $a_2 = a \oplus a_1$ הקידוד יהיה (a_1, a_2)ולכן A שולחת את a_2 ל B, כך ש $a = a_1 \oplus a_2$.
באופן דומה B עבור b בוחר זוג מקרי $b = b_1 \oplus b_2$ ושולח את b_1 ל A.
נניח שנרצה לחשב שער $\oplus$:
אזי

$$V = V_1^A \oplus V_2^B$$
$$W = W_1^A \oplus W_2^B$$

והחישוב:

$$V \oplus W = \left(V_1 \oplus W_1\right)^A + \left(V_2 \oplus W_2\right)^B$$

אם נרצה לחשב שער $\odot$ (כפל או AND):

$$\begin{aligned} V \odot W &= \left(V_1^A \oplus V_2^B\right) \odot \left(W_1^A \oplus W_2^B\right) \\ &= V_1^A \odot W_1^A \oplus \underbrace{V_1^A \odot W_2^B} \oplus \underbrace{W_1^A \odot V_2^B} \oplus V_2^B \odot W_2^B \end{aligned}$$

את הזוגות המסומנים נחשב בעזרת OT.
A בוחרת ביט מקרי r^A ומכניסה ל OT את הזוג

$$\left(\left(V_1^A \odot 0\right) \oplus r^A, \left(V_1^A \odot 1\right) \oplus r^A\right)$$

ו B מכניס ל OT את W_2^B.
נבחין כי לפי הגדרה זו B מקבל כפלט את $c^B = \left(V_1^A \odot W_2^B\right) \oplus r^A$.
באופן דומה A יכולה לקבל את $c^A = \left(W_1^A \odot V_2^B\right) \oplus r^B$ (כאשר r^B נבחר באקראי על ידי בוב)
(ויובל מעיר - בסימונים הנוכחיים לסאבסקריפט 1 או 2 אין משמעות, והם תמיד קונסיסטנטיים עם A ו B שמופיעים בסופרסקריפט).
ואז:

$$(V \odot W)^A = V_1^A \odot W_1^A \oplus r^A \oplus c^A$$
$$(V \odot W)^B = V_2^B \odot W_2^B \oplus r^B \oplus c^B$$

וכאשר B ירצה לקבל את התוצאה הסופית - הוא יכול לקבל מ A את המשלימים ה XORים הדרושים לשערי הפלט, ולחשב את התוצאה הסופית.
אם רוצים לעשות את זה עם שחקנים לא אמינים, דורשים שכל אחד מהם יתחייב על כל פעולה ונעבוד עם הוכחות אפס ידיעה.

7 הוכחות באפס ידיעה - $Zero\,Knowledge\,Proofs$

7.1 מערכת הוכחה אינטראקטיבית $(IP - Interactive\,Proofs$

הראינו כיצד ניתן לבצע חישוב פרטי בין שחקנים אמינים. נרצה לקבל פרוטוקולים בטוחים גם ללא ההנחה שהשחקנים אמינים. ניעזר באחד הכלים המשמעותיים שפותחו בתחום - הוכחות באפס ידיעה. תחילה נגדיר מהי מערכת הוכחה אינטראקטיבית.

הגדרה 7.1 תהי שפה $L \subseteq \Sigma^*$.
מערכת הוכחה אינטרקטיבית לשפה L היא פרוטוקול בין שני שחקנים $P(Prover), V(Verifier)$ כך ש:
1. V מכונה פולינומיאלית הסתברותית (ב-$|X|$).
2. P לא מוגבל חישובית אלא אם נציין אחרת בצורה מפורשת.
3. בהינתן $x \in \Sigma^*$, אם $x \in L$ אזי P משכנע את V בהסתברות 1 שאכן $x \in L$. באופן פורמאלי:

$$Pr[(P^*, V)[x] = 1)] = 1$$

ובהינתן $x \in \Sigma^*$ כך ש-$x \notin L$, לכל P^* (לאו דווקא מוכיח אמין) V לא משתכנע בהסתברות לכל הפחות $1/2$. באופן פורמאלי:

$$Pr[(P^*, V)[x] = 1)] \leq \frac{1}{2}$$

אם לשפה L קיימת מערכת הוכחה אינטראקטיבית, נאמר כי $L \in IP$ ($IP - Interactive\,Proof$).

ניתן דוגמא לשפה שלא יודע אם היא בNP (אך ידוע כי היא ב-$co - NP$) ואנו נראה עבורה הוכחה אינטראקטיבית.
תהי השפה $Non - Iso = L = \{(G_0, G_1) | G_0, G_1\, are\, non - isomorphic\, graphs\}$. ידוע כי השפה המשלימה $\bar{L} = Iso$ הינה ב-NP ומכאן $L \in co - NP$ אך לא ידוע האם $L \in NP$.

טענה 7.2 $Non - Iso \in IP$

הוכחה: נראה פרוטוקול שיענה לדרישות הגדרה 7.1. נניח כי מספר הקודקודים ב-2 הגרפים שווה ל-n (ברור כי אם מספר הקודקודים אינו שווה, הבעיה אינה מעניינת).
1. V בוחר ביט מקרי $b \in \{0, 1\}$ ופרמוטציה מקרית $\pi \in S_n$.
2. V מפעיל את H על G_b ושולח את $\pi(G_i) = H$ ל-P.
3. P מחזיר ל-V ביט $c \in \{0, 1\}$ (P מנסה להחזיר את האינדקס של הגרף האיזמורפי לגרף שקיבל).

4. V מקבל אם"ם $b = c$.

נשים לב כעת כי התנאים הדרושים מתקיימים. עבור $x = (G_0, G_1) \in Non - Iso$, H איזומורפי ל־G_b איך אינו איזומורפי ל־G_{1-b}. Pאינו מוגבל חישובית ולכן בהסתברות 1 יכול למצוא את G_b, כלומר את b ולשלוח ל־Vשמצידו ישתכנע. מנגד, עבור $x = (G_0, G_1) \notin Non - Iso$, H איזומורפי הן ל־G_0 והן ל־G_1. יתרה מכך, Hאינו תלוי ב־bמכיוון ש־π פרמוטציה מקרית. מכאן, ההסתברות ש־P^*מנחש נכונה את b היא בדיוק $1/2$. ■

הערה 7.3 אנו יודעים להגיד כי $IP = PSPACE$. לא ניכנס לזה בקורס.

7.2 $Zero\,Knowledge$

הערה 7.4 נשים לב כי ההגדרה שניתנה עבור מערכת הוכחה אינטראקטיבית 'מגנה' על המוודא V מפני ניסיון של P להוכיח לו הוכחה לא נכונה. לעיתים רבות, עולה המוטיבציה לגונן גם על P מפני היכולת של V לקבל מידע נוסף מלבד ההוכחה כשלעצמה בפרוטוקול התקשורת בינו לבין P. להוכחות אינטראקטיביות שמקיימות את התכונה הזו אנו קוראים הוכחות באפס ידיעה. נגדיר באופן פורמאלי.

הגדרה 7.5 נתבונן במערכת הוכחה אינטראקטיבית $(P,V)[x,y]$ כאשר ההוכחה מתייחסת לשאלה האם $x \in L$ ו־y הינו קלט נוסף. נאמר **שמערכת הוכחה אינטראקטיבית הינה באפס ידיעה** (סטטיסטית/חישובית) אם לכל V^* (לאו דווקא V הפועל לפי הפרוטוקול), קיים סימולטור S_{V^*} כך ש־S_{V^*} מכונה פולינומיאלית הסתברותית המקיימת:

$$\forall x \in L,\, y \in \Sigma^*, view_{V^*}[(P,V)[x,y]] = S_{V^*}(x,y)$$

במילים, y מתפקד על תקן $prior\,knowledge$ של V. ההגדרה מבטיחה כי V אינו יכול לנצל את הידע המוקדם שברשותו על מנת להפיק מהאינטרקציה עם P מידע נוסף. זאת מפני ש־S בהינתן x, y מסוגל להפיק את אותו המידע. השיוויון בהגדרה הינו שיוויון סטטיסטי/חישובי בין התפלגויות .אם לשפה L קיימת הוכחה באפס ידיעה אנו נסמן $L \in ZK$. נבדיל כאמור בין שפות עבורן קיימת הוכחה בחוסר ידיעה סטטיסטית (שם נסמן $L \in SZK$) לבין שפות עבורן קיימת הוכחה בחוסר ידיעה חישובית (שם נסמן $L \in CZK$). חשוב לשים לב כי הדרישה קיימת רק עבור מילים בשפה.

נשים לב כי ההוכחה האינטראקטיבית שהראינו עבור $Non - Iso$ אינה בחוסר ידיעה. לדוגמא, יתכן כי V מחזיק קלט נוסף $y = G$ כאשר V יודע כי G הינו גרף האיזומורפי ל־G_0 או ל־G_1 אך אינו יודע למי מהם. V האמיתי יפעל לפי הפרוטוקול ואז ההוכחה תהיה באפס ידיעה. אך V^* אחר יכול לשלוח את את G ל־P. P שאינו מוגבל חישובית ימצא לאיזה מבין הגרפים G_0, G_1 איזומורפי G ויחזיר את התשובה ל־V. ברור כי הוכחה שכזו אינה בחוסר ידיעה שכן V הפיק מן הפרוטוקול מידע נוסף מן ההוכחה (ניתן גם באופן פורמאלי להראות כי לכל סימולטור המוגבל חישובית אין יכולת להפיק את המידע הזה).

תרגיל: הראו פרוטוקול באפס ידיעה (סטטיסטי) ל־Non-Iso נראה כעת הוכחה באפס ידיעה (סטטיסטית) לבעית הגרף איזומורפיזם. נגדיר באופן פורמאלי $Iso = L = \{(G_0, G_1) | G_0, G_1\, are\, isomorphic\, Graphs\}$.

טענה 7.6 $Iso \in SZK$

כאמור, אנו נראה כעת הוכחה באפס ידיעה (סטטיסטית) לבעית הגרף איזומורפיזם:

1. P בוחר פרמוטציה מקרית π,מחשב את $H = \pi(g_0)$ ושולח ל־V את H.
2. V בוחר $b \in_R \{0,1\}$ ושולח ל־P.
3. תהי פרמוטציה $\phi \in S_n$ כך ש־$\phi(G_0) = G_1$ (ϕ קיימת אם"ם G_0, G_1 איזומורפיים). אם $b = 0$ אזי P שולח ל־Vאת $\psi = \pi^{-1}$. אחרת ($b = 1$), P שולח ל־V את $\psi = \phi \circ \pi^{-1}$.
4. Vבודק ומקבל אם"ם $\psi(H) = G_b$.

ראשית, נראה כי הפרוטוקול שהראינו אכן מהווה הוכחה ב־IP. עבור $x = (G_0, G_1) \in Iso$, P יכול לחשב את ϕואת ψ נכונה ולשכנע את V. עבור $x = (G_0, G_1) \notin Iso$ברור כי P^* יתפס בהסתברות חצי, שכן בהסתברות $1/2$ V יגריל $b = 1$ ואז P לא יצליח לייצר ψ מתאימה.

על מנת להוכיח כי ההוכחה היא באפס ידיעה נתאר סימולטור S(עבור v^* כלשהו): נבחר ביט מקרי $b' \in \{0,1\}$ ופרמוטציה מקרית $\psi \in S_n$. מחשבים $H = \psi^{-1}(G_b)$ ואז נותנים את H ל־V^*. V^* בוחר ביט $b \in \{0,1\}$. אם $b = b'$ פולט את (b', H, ψ) אחרת, נחזור על התהליך. כעת, נשים לב כי עבור $(G_0, G_1) \in Iso$ לא קיימת תלות בין H, b'. לכן b, b' אינן תלויים ומכאן בהסתברות $1/2$ מתקיים $b = b'$ ולכן בתוחלת נעצור אחרי פעמיים. בנוסף,

כשעצרנו ניתן להתבונן ולראות כי קיים שיוויון סטטיסטי בין ההתפלגות של (b', H, ψ) המתארת את התקשורת מצידו של V^* לבין זו המתארת את השיחה כפי שמייצר הסימולטור.

7.3 $Zero\,Knowledge\,Proofs\,Of\,Knowledge$

לעיתים אנו מעוניינים שה־$Prover$ יוכיח הוכחות ידיעה. כלומר, נרצה שמלבד ההוכחה, תינתן לנו הוכחה ש־P מכיר את העד להוכחה. לדוגמא, עבור השפה $L = Iso$ נרצה שהמוכיח יוכיח איזומורפיזם ונרצה גם להשתכנע שהעד מכיר את הפרמוטציה ϕ כך ש־$\phi(G_0) = G_1$. באופן כללי, נניח כי קיים יחס R פולינומיאלי (בדיקה פולינומיאלית) ועבור x נתון (P, V מכירים את x), P מכיר y כך ש־$R(x, y)$ ורוצה לשכנע את V בכך. המקרים המעניינים הם כאשר P גם כן מוגבל חישובית ושחוסר הידיעה של V הוא סטטיסטי/חישובי. באופן פורמאלי, רוצים מערכת הוכחה באפס ידיעה לשפה $L = \{x | \exists y\, s.t\ |y| \leq n^c \wedge R(x, y)\}$ (לדוגמא x יכול להיות זוג גרפים (G_0, G_1) ו־y יהיה פרמוטציה ϕ כך ש־$\phi(G_0) = G_1$). מצפים שאם P מוכיח אזי הוא יודע את y. נגדיר זאת.

הגדרה 7.7 נאמר שמערכת הוכחה היא **הוכחה של ידיעה** עם הסתברות שגיאה $\epsilon > 0$ אם לכל מוכיח P^*, לכל $\delta \in [0, 1 - \epsilon]$ ולכל x קיים אלגוריתם הסתברותי $K(P^*, x)$ כך שאם

$$Pr[P^*\ convinces\ V\ that\ x \in L] \geq \delta + \epsilon$$

אזי

$$Pr[K(P^*, x)\ computes\ y\ s.t\ R(x, y)] \geq \delta$$

וזמן הריצה של K פולינומיאלי ב־$|X|$ ובזמני הריצה של V, P^*.

טענה 7.8 הפרוטוקול שהצגנו ב־ZK לשפה ISO הוא הוכחת ידיעה ל־ISO (באפס ידיעה) עם הסתברות שגיאה $1/2$.

הוכחה: נשתמש באותם סימונים בהם השתמשנו בהוכחה טענה 7.6. נניח כי P^* מצליח בהסתברות $\leq \frac{1}{2} + \delta$. H נבחר לפי התפלגות כלשהי D. מתקיים, אם כך:

$$Pr_{H \sim D,\, b \in_R \{0,1\}}[H \cong G_b] \geq 1/2 + \delta$$

משתמע מכך שקיימת קבוצה של H־ים עם משקל δ כך ש־התשובה נכונה הן עבור $b = 0$ והן עבור $b = 1$. נתאר את K. בהינתן x (מקודד זוג גרפים (G_0, G_1)), מריצים את P^* על מנת לקבל את H. כעת, מריצים במקביל את P^* עם הדרישות $b = 0, b = 1$. כאמור, בהסתברות δ P^* ישכנע את V ב־2 ההרצות. במקרה זה נסיק $\phi = \psi_1 \circ \psi_0^{-1}$. כלומר, בהסתברות δ קיבלנו את העד ϕ לאיזומורפיזם בין G_0 ל־G_1. ■

הערה 7.9 אם נחזור לרגע לפרוטוקול שהראינו ב־IP ל־$Non - Iso$ ניזכר כי הוא לא ב־ZK כי V^* יכול לשלוח קלט אחר שנובע מ־$Prior\,Knowledge$ שיש לו וכך להפיק מידע נוסף. אם V יוכיח באפס ידיעה הוכחת ידיעה שהוא יודע את ההעתקה מ־G_0 או G_1 ל־H ששלח אזי בעיה זו תיפטר (כתרגיל - להראות זאת בפירוט).

7.4 $CZK\ Proofs\ For\ NP$

נתחיל במשפט המכרכזי של חלק זה.

משפט 7.10 אם קיימות פונקציות חד כיווניות אזי לכל שפה $L \in NP$ קיימת הוכחה באפס ידיעה חישובית (ההוכחה גם תהיה הוכחת ידיעה עבור L). במילים אחרות, אם קיימות פונקציות חד־כיווניות אזי $NP \subseteq CZK$.

הוכחה: מספיק להוכיח זאת עבור שפה $L \in NP-Complete$ שכן אם נצליח לעשות כן, אזי עבור כל שפה $L' \in NP$ נבצע את הרדוקציה לשפה L, נוכיח עבורה ונסיק בהתאם. נעשה זאת עבור $L = Graph\,Hamiltonicity$. נתאר את הפרוטוקול:

1. P בוחר פרמוטציה מקרית π, מחשב $H = \pi(G)$ ואת מטריצת השכנויות של H אותה נסמן ב־A. P מתחייב על π ועל כל איבר $A_{i,j}$ במטריצת השכנויות ע"י BC (כאן מסתתרת הנחנתנו שקיימות פונקציות חד כיווניות. נניח למשל שאנחנו משתמשים בפרוטוקול 5.2. עוד נשים לב כי ב־G קיים מעגל המילטוני אם"ם ב־H קיים כזה.)
2. V בוחר $b \in_R \{0,1\}$ ושולח ל־P.
3. אם $B = 0$, P פותח את כל התחייבויותיו. אם $b = 1$, P פותח את כל האיברים במטריצה המשתתפים במעגל המילטון (ב־H).
4. V בודק שההתחייבויות נפתחות כשורה. אם $b = 0$ אזי V אם הגרף שהתקבל מ־P הוא אכן $\pi(g)$. אם $b = 1$, V מקבל אם התתחייבויות נפתחו כשורה ואכן השביל המתואר הוא מעגל המילטוני.

תחילה נשתנכע כי הפרוטוקול ב־IP. נשים לב כי אם קיים מעגל המילטוני ב־G אזי P יתחייב על H שגם בו קיים מעגל המילטוני. P יוכל להתחייב כהלכה ולפתוח את התחייבויותיו כהלכה במקרה $b = 0$ ויוכל גם להציג מעגל המילטוני ב־H במקרה $b = 1$. מכאן, V ישתכנע במקרה זה בהסתברות 1. אם לא קיים מעגל המילטוני ב־, אזי, אם P^* לא התחייב בשלב ה־BC על עותק איזומופרפי ל־G אזי בהסתברות $1/2$ V יגריל $b = 0$ ו־P^* יתפס. אם P^* מתחייב כהלכה, אזי בהסתברות $1/2$ יוגרל $b = 1$ ו־P לא יוכל להציג מעגל המילטוני ב־H.

כעת, על מנת להראות כי הפרוטוקול ב־CZK, נציג סימולטור S:

1. S בוחר $b' \in_R \{0,1\}$. אם $b' = 0$, S מגריל פרמוטציה π ומתחייב עליה ועל מטריצת השכנויות של $H = \pi(G)$. אם $b' = 1$, S מחליף את G במעגל G' מקרי ומתחייב על G'.
2. S מריץ את V^* שמצידו משיב עם ביט $b \in \{0,1\}$. אם $b = b'$ ממשיכים. אחרת, מאתחלים.

ראשית, נטען כי ההסתברות למאורע $b = b'$ הינה $1/2 - \epsilon$ כאשר ϵ הינה ההסתברות שמכונה פולינומיאלית כ־V^* יכולה להבדיל בין ההתפלגויות במקרים $b' = 0$ ו־$b' = 1$. לכן, תוחלת מספר האיטרציות נמוכה ($\frac{1}{\frac{1}{2}-\epsilon}$). נבחן את פני הדברים כאשר $b = b'$. אם $b = 0$, אזי ההתפלגויות זהות. אם $b = 1$, קיים שוני בין ההתפלגויות אבל יכולתו של V^* להבחין חישובית בין ההתפלגיות מוגבלת ע"י הפרמטרים של ה־BC. ■