עבודה מספר 1 באלגוריתמים מתקדמים:

שאלה מספר 1:

צריך להוכיח את קיומו של אלגוריתם רנדומלי שגיאה אפסי.

באופן כללי כל אלגוריתם שראינו עד עכשיו, תמיד הייתה שגיאה כלשהי לתשובה לא נכונה, אבל אפשר להוריד את זה על ידי מספר רב של חזרות.

קיים ZPP סט של שפות המכילות את כל השפות שיש אלגוריתם רנדומלי שגיאה-אפסי שמכריע אותם ורץ בזמן פולינומי.

קיים BPP על כל הבעיות A שיש להן אלגוריתם אקראי יעיל.

נניח $RP = BPP\left(0\ ,\frac{1}{2}\right)$ וגם $coRP = BPP\left(\frac{1}{2}, 1\right)$

כלומר עבור RP :

$$if\ x \in A, \Pr[M\ accepts] \geq \frac{1}{2}$$
$$if\ x \notin A, \Pr[M\ accepts] \leq 0$$

ועבור $coRP$:

$$if\ x \in A, \Pr[M\ accepts] \geq 1$$
$$if\ x \notin A, \Pr[M\ accepts] \leq \frac{1}{2}$$

כדי להקטין את השגיאה נרצה שמה שבשפה יהיה קרוב ל1 ומה שלא בשפה יהיה קרוב ל0.

a. צריך להראות שניתן להגדיר באופן שווה את ZPP כמחלקה של שפות A עבורה יש אלגוריתם M אקראי עם המאפיינים הבאים:

1. זמן הריצה של אלגוריתם M הוא פולינומי במקרה הגרוע לכל קלט x.
2. בהינתן קלט x , האלגוריתם M מחזיר אם x נמצא בשפה A או אומר אני לא יודע.
3. אם M לא אומר "לא יודע", התשובה שלו תמיד נכונה.
4. אם M אומר " אני לא יודע" ההסתברות של רוב היא $\frac{1}{3}$ (ההסתברות היא מעל המטבעות האקראיים של M ולא מעל הקלטים).

פתרון :

כדי להראות שניתן להגדיר את ZPP באופן שווה כמחלקה של שפות A המקיימות את המאפיינים 1,2,3,4 אדגים את שני הכיוונים:

כל שפה ZPP עומדת במאפיינים 1,2,3,4

וגם כל שפה שמספקת את המאפיינים 1,2,3,4 נמצאת בZPP.

<u>נתחיל מחלק ראשון: נוכיח שכל שפה שנמצאת בZPP עומדת במאפיינים 1,2,3,4</u>

נניח שA היא שפה בZPP, כלומר קיים אלגוריתם הסתברותי בזמן פולינומי M שמכריע את A בהסתברות שגיאה של $\frac{1}{3}$ לכל היותר.

1. מכיוון ש A נמצא בZPP, קיים אלגוריתם הסתברותי בזמן פולינומי M שמחליט על A. ולכן זמן הריצה של M הוא פולינומי במקרה הגרוע ביותר עבור כל קלט x .
2. לפי ההגדרה של ZPP, לM יש הסתברות לשגיאה של לכל היותר $\frac{1}{3}$. זה אומר שלכל קלט x , אלגוריתם M מוציא:

* כן אם הקלט נמצא בשפה A
* לא אם x אינו בשפה A
*"אני לא יודע"

לכן בהינתן קלט x, M עונה אם נמצא בA או אומר אני לא יודע.

3. מכיוון שM הוא אלגוריתם הסתברותי בזמן פולינומי שמכריע את שפה A עם הסתברות שגיאה לכל היותר של שליש, זה אומר שבכל פעם שM מוציא את האופציות : "כן" או "לא" (כלומר לא אומר אני לא יודע) התשובה שלו תמיד נכונה.
4. לפי ההגדרה שלZPP , לM יש הסתברות לשגיאה של לכל היותר $\frac{1}{3}$. זה אומר שההסתברות שM יגיד "אני לא יודע" היא לכל היותר $\frac{1}{3}$.

לפיכך כל שפה בZPP עומדת במאפיינים 1,2,3,4.

<u>נמשיך לחלק שני: כל שפה שמספקת את המאפיינים 1,2,3,4 נמצאת בZPP</u>

נניח שקיימת שפה A שעונה על המאפיינים 1,2,3,4 כמו שמתואר בשאלה. אנחנו צריכים להראות ש A נמצא בZPP. לשם כך נבנה אלגוריתם מונטה קרלו ב זמן פולינומי $\grave{M}$ שמכריע את A עם הסתברות שגיאה של לכל היותר $\frac{1}{3}$.

אלגוריתם $\grave{M}$:

- הפעל את M על קלט x בזמן פולינומי.
- אם M מוציא "כן" או "לא" תכתוב את התשובה ותעצור.
- אם M מוציא "אני לא יודע", חזור על שלב הראשון לכל היותר שלוש פעמים.
- אם M מוציא "כן או "לא" בכל אחת מהחזרות האלה, תדפיס את אותה תשובה והפסק. אחרת תדפיס "אני לא יודע".

כעת ננתח את המאפיינים של $\grave{M}$:

1. מכיוון ש$\grave{M}$ מריץ את M שיש לו זמן ריצה פולינומי ומבצע מספר פולינומי של חזרות במידת הצורך, זמן הריצה של $\grave{M}$ הוא פולינומי במקרה הגרוע ביותר בכל קלט x
2. $\grave{M}$ מפעיל את M שמוציא " כן " או "לא" או "לא יודע".
   אם M מוציא "כן" או "לא", $\grave{M}$ מספק אותה תשובה.
   אחרת, $\grave{M}$ חוזר על התהליך עד שלוש פעמים . ואם מוציא " כן " או "לא" בכל אחת מהחזרות, $\grave{M}$ מוציא אותה תשובה.
   אם M לא מוציא תשובה סופית בתוך החזרות, $\grave{M}$ עונה "אני לא יודע."
   ובכך השגנו את זה שבהינתן קלט x, M עונה אם x נמצא בA או אומר "אני לא יודע".
3. מכיוון שM עונה על תכונה זו, פירוש הדבר שכאשר M מוציא " כן " או "לא" , התשובה שלו תמיד נכונה. אלגוריתם $\grave{M}$ פשוט מעביר את הפלט שלM, כך שאם M לא אומר "אני לא יודע", $\grave{M}$ יספק אותה תשובה נכונה. לפיכך אם M לא אומר "אני לא יודע", התשובה שלו תמיד נכונה.
4. בבניה של $\grave{M}$ אנו חוזרים על התהליך עד שלוש פעמים אם M מוציא " אני לא יודע". אם $\grave{M}$ מגיע למספר המקסימלי של חזרות וM עדין לא פלט "כן" או "לא", רק אז $\grave{M}$ אומר "לא יודע". מכיוון שM אומר "לא יודע" בהסתברות של שליש לכל היותר, יוצא שגם $\grave{M}$ אומר "אני לא יודע" בהסתברות של לכל היותר $\frac{1}{3}$.

לפי המאפיינים 1,2,3,4 הראינו שהשפה A עומדת בתנאים לZPP. לפיכך כל שפה שמספקת את המאפיינים של 1,2,3,4 היא אכן בZPP.

הראינו את שני כיווני השקילות, ולכן הראינו שניתן להגדיר את ZPP באופן שווה כמחלקה של שפות A המקיימות את המאפיינים המתוארים בשאלה.

b. צריך להראות ש $RP \subseteq BPP$ *and similarly* $coRP \subseteq BPP$.

ידוע לנו כי:

נניח $RP = BPP\left(0\ ,\frac{1}{2}\right)$ וגם $coRP = BPP\left(\frac{1}{2},1\right)$

כלומר עבור RP :

$$if\ x \in A, \Pr[M\ accepts] \geq \frac{1}{2}$$
$$if\ x \notin A, \Pr[M\ accepts] \leq 0$$

ועבור $coRP$:

$$if\ x \in A, \Pr[M\ accepts] \geq 1$$
$$if\ x \notin A, \Pr[M\ accepts] \leq \frac{1}{2}$$

RP היא מחלקת השפות שעבורו קיים אלגוריתם אקראי שפועל בזמן פולינומי. יש לו $\frac{1}{2}$ הסתברות לשגיאה מקסימלית. והוא מוציא "כן" אם הקלט בשפה או מוציא "לא" או "לא יודע" אם הקלט אינו בשפה.

באופן דומה $coRP$ היא מחלקת השפות שעבורן קיים אלגוריתם אקראי הפועל בזמן פולינומי בעל $\frac{1}{2}$ כהסתברות השגיאה המקסימלית ומוציא "לא" אם הקלט אינו בשפה ואחד מהם מוציא "כן" או "לא יודע" אם הקלט בשפה.

אזBPP היא מחלקת השפות שעבורן קיים אלגוריתם אקראי הפועל בזמן פולינומי, בעל הסתברות שגיאה מוגבלת ומוציא את התשובה הנכונה בהסתברות גבוהה. אנחנו רוצים להראות שRPו$coRP$ הן שתי קבוצות משנה של BPP.

<u>נתחיל עם הוכחה ש$RP \subseteq BPP$</u>

נניח שA היא שפה בRP, ניקח בחשבון את האלגוריתם האקראי M שמכריע את השפה A. M פועל בזמן פולינומי ויש לו הסתברות מרבית לשגיאה של $\frac{1}{2}$. כדי להמיר את M לאלגוריתם BPP אנחנו יכולים לשנות אותו באופן הבא:

אלגוריתם $\grave{M}$ (גרסת BPP שלM):

- הפעל את M על קלט x עבור מספר פולינום של חזרות, בכל פעם באופן עצמאי.
- אם M מוציא "כן" החזר "כן"
- אם M מוציא "לא" בכל החזרות, החזר "לא"
- אחרת החזר " אני לא יודע"

מכיוון שM יש הסתברות מקסימלית לשגיאה של חצי, על ידי הפעלתו מס]ר פעמים, ההסתברות לקבל תשובה לא נכונה יורדת בצורה אקספוננציאלי. לכן ההסתברות לשגיאה של $\grave{M}$ יכולה להיעשות קטנה באופן שרירותי על ידי הגדלת מספר החזרות, תוך שמירה על זמן ריצה פולינומי. לפיכך A נמצאת בBPP.

באופן דומה נראה ש $coRP \subseteq BPP$ על ידי ביצוע דומה.
ולכן הוכחנו ש $RP \subseteq BPP \text{ and similarly } coRP \subseteq BPP$.

c. צריך להוכיח ש$ZPP = RP \cap coRP$.

כדי להוכיח את השוויון הזה צריך להראות את שני הכיוונים.

- $ZPP \subseteq RP \cap coRP$: כל שפה ב-$ZPP$ נמצאת גם ב-RP וב-$coRP$.
- $RP \cap coRP \subseteq ZPP$: כל שפה בהצטלבות של RP ו-$coRP$ נמצאת גם היא ב-ZPP.

<u>נתחיל מהחלק הראשון $ZPP \subseteq RP \cap coRP$:</u>

נניח שA היא שפה בZPP, כלומר קיים אלגוריתם אקראי אפס-שגיאות M שמכריע את A וזמן הריצה הצפוי שלו הוא פולינומי.
מכיוון שלM יש אפס-שגיאה זה אומר שM הוא גם אלגוריתם בRP שהרי RP מאפשר הסתברות של חצי לשגיאה.
בנוסף, מכיוון שלM יש זמן ריצה פולינומי, הוא עונה על התנאים להיות בBPP.
לכן A נמצא בRP וA נמצא בBPP.
באו]ן דומה A נמצא ב$coRP$ וA נמצא בBPP.
לפיכך A נמצא בצומת של RP ו$coRP$.

<u>חלק שני $RP \cap coRP \subseteq ZPP$:</u>

נניח שA היא שפה בצומת של RP ו$coRP$, כלומר קיים אלגוריתם אקראי M שמכריע את A ושייך גם לRP וגם ל$coRP$.
מכיוון שM לא נמצא בRP. זה אומר שלM יש אפס-שגיאה וזמן ריצה פולינומי. באופן דומה מכיוון שM נמצא ב$coRP$, יש לו אפס שגיאה וזמן ריצה פולינומי. לכן M עונה על המאפיינים של אלגוריתם אקראי עם שגיאות אפס שמכריע את A. ויש לו זמן ריצה פולינומי.. לפיכך A נמצא בZPP

בשילוב שני החלקים, הראינו ש-$ZPP \subseteq RP \cap coRP$ ו-$RP \cap coRP \subseteq ZPP$.
לכן, אנו מסיקים ש-ZPP שווה לצומת של RP ו-$coRP$, כלומר $ZPP = RP \cap coRP$.

<u>שאלה מספר 2:</u>

צריך להוכיח את הלמה של שוורץ זיפל.

בהינתן $f(x_1,..,x_n)$ פולינום שונה מאפס כאשר d המעלה של הפולינום עם מקדמים רציונליים. נתון $S \subseteq \mathbb{Q}$ כאשר S תת קבוצה של מספרים רציונליים ו$|S|$ זה גודל של תת קבוצה אז נטען ש: $Pr_{a_1,..,a_n \subseteq S}[f(a_1,..,a_n)=0] \leq \frac{d}{|S|}$ כאשר נבחר את $a_1,..,a_n$ בהתפלגות רנדומלית ואחידה מ S.

באופן כללי הלמה מספקת אלגוריתם הסתברותי לבדיקה האם פולינום רב-משתני הוא אפס.

<u>הוכחה באינדוקציה:</u>

כאשר n=1 נתון $f(x)$ הוא פולינום עם <u>משתנה יחיד</u> בדרגה d. מכיוון שלפולינום שאינו אפס בדרגה d יש לכל היותר d שורשים בכל תחום, אפשר להסיק שהלמה נכונה. $Pr_{a_1,..,a_n \subseteq S}[f(a_1,..,a_n)=0] \leq \frac{d}{|1|} = d$

נניח שהלמה נכונה עבור n-1 משתנים.

נתון $f(x_1,..,x_n)$ פולינום שונה מאפס כאשר d המעלה של הפולינום עם מקדמים רציונליים. נניח ש $f(x_1,x_2,..,x_{n-1},x_n) = g(x_1,x_2,..,x_{n-1})$ אז $g(x_1,x_2,..,x_{n-1})$ הוא פולינום שאינו אפס בעל דרגה d במשתנים $x_1,..,x_{n-1}$. לפי השערת האינדוקציה ההסתברות ש$g(a_1,..,a_{n-1})=0$ עבור המשתנים $a_1,..,a_{n-1}$ שנבחר באקראי מS היא לכל היותר $\frac{d}{|S|}$.

נסתכל על הפולינום $f(x_1,..,x_n)$. לפי המקרה עם של הלמה עם משתנה יחיד, ההסתברות ש $f(a_1,..,a_n)=0$ עבור $a_1,..,a_n$ שנבחר באקראי מS היא לכל היותר $\frac{d}{|S|}$. אלא אם כן ש$g(a_1,..,a_{n-1})$ זהה לאפס.

אם ש$g(a_1,..,a_{n-1})$ הוא פולינום שהוא אפס, אז $f(x_1,..,x_n)$ הוא פולינום בעל דרגה לכל היותר d-1 בx והלמה נובעת מהמקרה עם משתנה יחיד.

אחרת, $g(a_1,..,a_{n-1})$ אינו אפס, ולכן ל$f(x_1,..,x_n)$ יש לכל היותר d שורשים.

ומכאן ההסתברות ש-$f(a_1,..,a_n)=0$ עבור $a_1,..,a_n$ שנבחר באקראי מ-S היא לכל היותר |d/|S.

לכן, הלמה של שוורץ-זיפל מתקיימת עבור כל פולינום שאינו אפס בדרגה d ב-n משתנים.

<u>שאלה מספר 3:</u>

צריך להוכיח את שאם G הוא גרף מחובר (קשיר) אז $\lambda_2 \leq 1 - \frac{1}{4n^4}$

נתון גרף G די-רגולרי וA מטריצת השכנויות המנורמלת שלו. הגדרנו בביתה ש$B = \frac{1}{2}I + \frac{1}{2}A$. וגם שיש לנו וקטור עצמי של B בסדר יורד כך ש $0 \leq \lambda_i \leq \lambda_{i-1} \leq \cdots \leq \lambda_1 = 1$.

צריך להוכיח את שאם G הוא גרף מחובר (קשיר) אז $\lambda_1 \leq 1 - \frac{1}{4n^4}$

פתרון:

א. תן ל x להיות וקטור עצמי מנורמל המתאים לλ_2 כך ש $||x||_2 = 1$. צריך להראות שיש קורדינטה i כך ש $|x| \geq \frac{1}{\sqrt{n}}$

<u>הוכחה:</u>

יהי x וקטור עצמי השייך לערך עצמי λ_2 המקיים $||x||_2 = 1$.

לפי הגדרת הנורמה נקבל עבור x כך : $\sum_{k=1}^{n} x_k^2 = 1$.

נניח בשלילה כי לכל קורדינטה מתקיים $|x| < \frac{1}{\sqrt{n}}$ ואם נחשב את הנורמה שלו נקבל –

$$\sum_{k=1}^{n} x_k^2 < \sum_{k=1}^{n} |\frac{1}{\sqrt{n}}|^2 = \sum_{k=1}^{n} \frac{1}{n} = 1$$

וזאת סתירה לכך ש$\sum_{k=1}^{n} x_k^2 = 1$ כמו שציינו. ולכן המשפט לא מתקיים וכן קיימת קורדינטה $|x_i| \geq \frac{1}{\sqrt{n}}$

ב. נניח ללא אובדן הכלליות שעבור הקורדינטה i מתקיים $x_i \geq \frac{1}{\sqrt{n}}$

הראה שיש קורדינטה j שונה כך ש$x_j \leq 0$.

רמז: השתמש בעובדה ש1 הוא וקטור עצמי מתאים ל λ_1 וכי וקטורים עצמיים התואמים לערכים עצמיים הם אורתוגונליים.

<u>הוכחה:</u>

בשיעור הוכחנו ש $\lambda_1 = \mu_1 = 1$ כלומר הערך העצמי הוא 1 והוא מתאים לווקטור עצמי שכולו 1 : $v_1 = \frac{1}{\sqrt{n}}$

(מנורמל ב1 חלקי שורש n כדי שסכום הריבועים יהיה 1) והוא שייך למטריצה B.

בשיעור אמרנו שווקטורים עצמיים במטריצה B אורתונורמליים אחד לשני ולכן מתקיים : $< 1, x > = 0$

$$< 1, x > = \sum_{k=1}^{n} x_k \;= 0$$

נניח בשלילה שלא קיימת קורדינטה j המקיימת ש$x_j \leq 0$ כלומר לכל j מתקיים ש$x_j > 0$. לכן :

$$0 = < 1, x > = \sum_{k=1}^{n} x_k \;= \sum_{k \neq 1}^{n} x_k \;+ x_i \;\; \geq \frac{1}{\sqrt{n}}$$

וזו כמובן סתירה ולכן אפשר להגיד שכן קיימת קורדינטה j המקיימת $x_j \leq 0$.

ג. צריך להראות שקיימת קשת (u,v) בגרף G שעבורה מתקיים $|x_u - x_v| \geq \frac{1}{n\sqrt{n}}$

רמז: השתמש בהנחה שG הוא קשיר ומחפש מסלול מx_i לx_j.

<u>הוכחה:</u>

נסתכל על הביטוי שעבורו צריך להוכיח $|x_u - x_v|$ ננסה לפתוח אותו.
הגרף קשיר G ויש מסלול בין קודקוד I לj לכן אפשר להסתכל על המסלול מקודקוד i לקודקוד j.
נסמן את הקודקודים שעוברים דרכם במסלול: $i_1, i_2, \dots, i_{k-1}$.
בסעיפים הקודמים הוכחנו ש $|x_i| \geq \frac{1}{\sqrt{n}}$ וגם $0 \geq x_j$.
ועכשיו ניתן להסיק את הדבר הבא:

$$\frac{1}{\sqrt{n}} \leq |x_i - x_j| = \left|x_i - x_{i_1} + x_{i_1} - \cdots - x_{i_{k-1}} + x_{i_{k-1}} - x_{i_{k=j}}\right|$$
$$= \left|(x_i - x_{i_1}) + (x_{i_1} - \cdots - x_{i_{k-1}}) + (x_{i_{k-1}} - x_{i_{k=j}})\right|$$

כלומר שהמסלול הכי ארוך שיכול להיות בין i לj יהיה לכל היותר n. כיוון ויש n מחוברים מהסוג $(x_u - x_v)$.
הסכום הזה חסום מלמטה על ידי $\frac{1}{\sqrt{n}}$ לכן אם הוא סכום של n מחוברים לכל היותר, חייב להיות לפחות מחובר אחד שגדול מ $\frac{1}{n\sqrt{n}}$ (אחרת זה יהיה קטן מ $\frac{1}{\sqrt{n}}$)ומכאן קיים ש $|x_u - x_v| \geq \frac{1}{n\sqrt{n}}$
משל.

ד. צריך להצדיק את המעברים הבאים:

$$\lambda_2 = \lambda_2 \cdot \langle x, x\rangle = \langle \lambda_2 x, x\rangle = \langle Bx, x\rangle = \sum_{k,\ell} B_{k,\ell} \cdot x_k x_\ell$$
$$= \frac{1}{2}\sum_{k,\ell} B_{k,\ell}\left[x_k^2 + x_\ell^2 - (x_k - x_\ell)^2\right]$$
$$= \sum_k x_k^2 - \frac{1}{2}\sum_{k,\ell} B_{k,\ell}\cdot(x_k - x_\ell)^2 = 1 - \frac{1}{2}\sum_{k,\ell} B_{k,\ell}\cdot(x_k - x_\ell)^2$$

הצדקות:

- $\lambda_2 = \lambda_2 * \langle x, x\rangle$ ידוע שx הוא ערך עצמי אורתונורמלי והנורמה שלו היא 1.
- $\lambda_2 * \langle x, x\rangle = \langle \lambda_2 x, x\rangle$ מכפלת סקלר ונורמה
- $\langle \lambda_2 x, x\rangle = \langle Bx, x\rangle$ הגדרות של ערך עצמי ווקטור עצמי
- $\langle Bx, x\rangle = \sum_{k,l} B_{k,l} * x_k x_l$כפל מטריצות וחישוב נורמה
- $\sum_{k,l} B_{k,l} * x_k x_l = \frac{1}{2}\sum_{k,l} B_{k,l}[x_k^2 + x_l^2 - (x_k - x_l)^2]$ לפי נוסחאות כפל ידועות ובמקרה הזה $2ab = a^2 + b^2 - (a-b)^2$
- $\sum_{k,l} B_{k,l} * x_k x_l = \frac{1}{2}\sum_{k,l} B_{k,l}[x_k^2 + x_l^2 - (x_k - x_l)^2] = \sum_k x_k^2 - \frac{1}{2}\sum_{k,l} B_{k,l}(x_k - x_l)^2$ הוצאת גורמים החוצה
- $\sum_k x_k^2 - \frac{1}{2}\sum_{k,l} B_{k,l}(x_k - x_l)^2 = 1 - \frac{1}{2}\sum_{k,l} B_{k,l}(x_k - x_l)^2$ x הוא וקטור עצמי אורתונורמלי.

סיכום משפט:
קיימת קשת (u,v) ולכן $A_{u,v} = \frac{1}{d}$ כך ש $u \neq v$ ולכן מהגדרת $B = \frac{1}{2}I + \frac{1}{2}A$ מתקיימים: $B_{u,v} = \frac{1}{2}A_{u,v}$ ומהחסם שגילינו קודם מתקיים $B_{u,v} * (x_u - x_v)^2 \geq \frac{1}{2d} * \frac{1}{n^3}$
מההגדרות שראינו $B_{k,l} \geq 0$ וגם $(x_k - x_l)^2 \geq 0$ ולכן $\sum_{k,l} B_{k,l}(x_k - x_l)^2 \geq B_{u,v} * (x_u - x_v)^2$
ועכשיו אפשר לסכם שנקבל $1 - \frac{1}{2}\sum_{k,l} B_{k,l}(x_k - x_l)^2 \geq 1 - \frac{1}{4dn^3}$

<u>שאלה מספר 4:</u>

בבעיית השוויון, כל אחד מהשחקנים, אליס ובוב מקבל בקלט string עם n ביטים.
כל שחקן רואה את הקלט שלו אבל לא את הקלט של השחקן האחר.
אליס ובוב היו רוצים לברר האם הקלטים שלהם שווים או לא שווים.
אפשר להראות את זה שאם אליס ובוב דטרמיניסטיים אז משימה זו דורשת n ביטים לפחות של החלפות. לדוגמה- אליס יכולה לשלוח לבוב את כל המחרוזת שלה.
תכנן פרוטוקול אקראי שמאפשר להם לגלות האם הקלטים שלהם שווים או לא ודורשת שליחת סיביות של $O(logn)$ במקרה הגרוע.
ההסתברות לשגיאה של הפרוטוקול צריכה להיות לכל היותר שליש (כלומר, ההתנהגות של אליס ובוב עשויה להיות אקראית ועבור כל צמד כניסות, השחקנים טועים בהסתברות של לכל היותר שליש).
רמז: דרך אפשרית לעשות זאת היא לפרש את הקלטים כמספרים שלמים בינאריים בגודל של לפחות $2^n = N$ וכדי לחשב את יתרת המודולו p שלהם עבור p ראשוני אקראי שנבחר מתוך מרווח מסוים מתאים . אפשר להשתמש במאפיינים של מספרים ראשונים שהדגמנו בכיתה. תחשוב כמה מחלקים ראשונים ברורים יכולים להיות למספר של גודל לכל היותר n).

<u>פתרון:</u>

כדי לתכנן פרוטוקול אקראי המאפשר לאליס ובוב לקבוע האם הקלט שלהם שווה או לא ודורש שליחת סיביות $O(logn)$ רק במקרה הגרוע, נוכל לעשות שימוש במאפיינים של מספרים ראשונים.

<u>פרוטוקול אפשרי:</u>

- אליס ובוב מתרגמים את המחרוזות הבינאריות למספרים טבעיים שלמים בגודל לכל היותר $2^n = N$
  כאשר המחרוזת של אליס שמתקבלת היא: $a = \sum_{i\in[n]} a_i 2^{n-1}$
  המחרוזת של בוב שמתקבלת היא: $b = \sum_{i\in[n]} b_i 2^{n-1}$.

- אליס ובוב בוחרים מספר אקראי ראשוני במרווח מתאים. נקרא למספר הזה p כאשר
  $p \le \tau = tn \log tn$

- אליס מחשבת את המספר שלה מודולו p המספר הראשוני בצורה הזו: $F_p(a) = a(mod\ p)$ ומעבירה לבוב את המספר שהתקבל $F_p(a)$ וגם את p

- בוב מחשב את $F_p(b) = b(mod\ p)$ ומחזיר "שווה" אם $F_p(b) = F_p(a)$, אחרת "לא שווה".

<u>נכונות:</u>

אם $a = b$ אז $F_p(b) = F_p(a)$
הפרוטוקול מבטיח שאם הקלטים של אליס ובוב שווים, הם יגיעו תמיד לאותה מסקנה.
אם $a \neq b$ וגם $F_p(b) \neq F_p(a)$, אלא אם כן p מתחלק בצורה כזו: $-2^n < a - b < 2^n$
עם זאת, קיימת אפשרות לשגיאה אם הקלטים שלהם אינם שווים כלומר אם $a \neq b$ ו $F_p(b) = F_p(a)$
ניתן לתחום את הסתברות השגיאה באמצעות מאפיינים של מספרים ראשונים שלמדנו בכיתה.
ידוע שיש לכל היותר n ראשוני מובהק המחלק מספר קטן מ2^n.

עבור כל τ, מספר הראשוני הקטן מτ הוא $\prod(\tau)\sim\frac{\tau}{ln\tau}$.

ואז אפשר לטעון שאם $a\neq b$ אז $\mathbb{P}[F_p(b)=F_p(a)]\leq\frac{n}{\prod(\tau)}=O(\frac{nlog(tn)}{tnlog(tn)})=O(\frac{1}{t})$ וזה בעצם אומר שההסתברות לטעות – שזה במידה והמספר האקראי *p* מחלק את המחרוזות אבל המחרוזות לא שוות, הוא $\frac{1}{t}$. נתון לנו שהשגיאה צריכה להיות לכל היותר $\frac{1}{3}$ ומכאן ש*t=3*.