The Clean Paper

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

הוכחה קריפטוגרפית יכולה להסתיר את הסוד שלה כשקשה להוכיח שאין סימולטור

הוכחות אפס־ידע מאפשרות להוכיח טענה בלי לחשוף את העד. התיאוריה הקלאסית אומרת שאי אפשר להשיג זאת, במובן המלא, באמצעות הודעה אחת, ללא הכנה מהימנה ועם נאותות מושלמת. המאמר של *Ilango Rahul* אינו מעלים את אי־האפשרות הזאת. הוא משנה את היעד. במקום לדרוש שסימולטור באמת יתקיים, הוא דורש שמערכת ההוכחה שנבחרה לא תוכל להוכיח ביעילות שאין סימולטור. תחת הנחות מרכזיות במורכבות הוכחות ובקריפטוגרפיה, די בכך כדי לשחזר, תכונה אחר תכונה, את ההשלכות ברות־ההפרכה ומבוססות־המשחק של אפס־ידע. זה אינו פרימיטיב אינטרנטי מוכן לשימוש. זו דרך תורת־הוכחתית להפוך אי־יכולת מתמטית להוכיח לכיסוי קריפטוגרפי.

Written by Cinzia Vaglio · Cross-checked by Vera Rubin · Edited by Michele Renda · Figures by Nova Vela · AI-assisted, human-reviewed

Paper: *Gödel in Cryptography: Effectively Zero-Knowledge Proofs for NP with No Interaction, No Setup, and Perfect Soundness*, Rahul Ilango, FOCS 2025 / IACR ePrint 2025/1296 · DOI: 10.1109/FOCS63196.2025.00058

הטריק אינו להוכיח שהסוד מוסתר

נתחיל בגרסה הפשוטה ביותר של אפס־ידע (zero-knowledge).

אליס רוצה לשכנע את בוב שלחידת סודוקו יש פתרון. אם היא שולחת את הפתרון, בוב משתכנע — אבל החידה נהרסת. היא רוצה משהו מוזר יותר: הוכחה לכך שקיים פתרון, בלי לחשוף את הפתרון עצמו.

זו ההבטחה של הוכחת אפס־ידע. המוכיח (אליס) משכנע את המאמת (בוב) שטענה מסוימת נכונה, בלי לחשוף דבר מעבר לעצם נכונותה של הטענה.

הבעיה היא שלהבטחה הזאת יש מחיר. להוכחה מתמטית רגילה יש שתי תכונות נוחות. היא **הודעה אחת**: כותבים אותה, מוסרים אותה והולכים. והיא בעלת **נאותות מושלמת** (perfect soundness): לטענה שקרית אין שום הוכחה תקפה. תוצאות אי־אפשרות קלאסיות אומרות שאפס־ידע חייב לוותר על שתי התכונות האלה — ולא רק על השילוב ביניהן; כל אחת מהן בפני עצמה אסורה.

ראשית, הוכחת אפס־ידע זקוקה לשיחה. אם אליס שולחת הודעה יחידה, בלי שלב הכנה מהימן (trusted setup) שנקבע מראש, הבטחת אפס־הידע קורסת — וזה נכון בלי קשר לכמה נאותות תהיו מוכנים להקריב בתמורה.

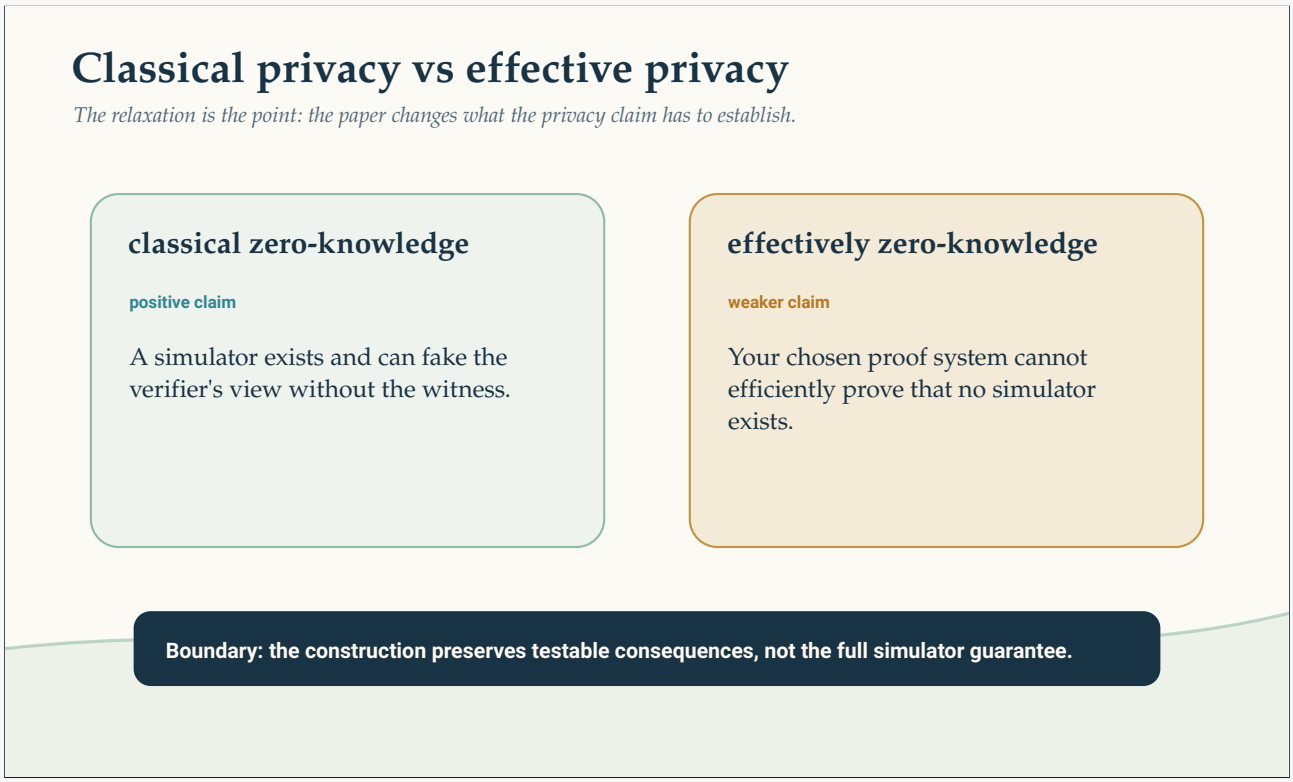
שנית, הוכחת אפס־ידע זקוקה לסבילות קטנה לטעות. מתברר שדרישה לנאותות מושלמת מחסלת בשקט גם את האינטראקציה: מאמת שאי אפשר להטעות אותו לעולם, בלי קשר לבחירות האקראיות שלו, יכול באותה מידה לקבע את הבחירות האלה מראש — וברגע שהמאמת צפוי, אליס יכולה לענות על הכול בהודעה אחת, בדיוק המקרה שכבר נשבר.

המאמר של Ilango Rahul עוסק בדרך לעקוף את הקיר הכפול הזה. לא באמצעות העמדת פנים שהקיר אינו קיים, ולא באמצעות יצירת אפס־ידע קלאסי בסביבה שבה הדבר בלתי אפשרי. המהלך עדין יותר: להחליש את משמעותו של "לא חושף דבר", אבל לעשות זאת באופן ששומר על תכונות האבטחה שקריפטוגרפים באמת יכולים לבדוק.

לתוצאה קוראים **אפס־ידע אפקטיבי** (effectively zero-knowledge).

אפס־ידע נחסם בשלוש דלתות — אינטראקציה, הכנה מהימנה ונאותות שאינה מושלמת. הבנייה של Ilango חומקת דרך דלת אחרת: ספר הכללים אינו מסוגל להפריך ביעילות את קיומו של הסימולטור.

Original diagram — The Clean Paper CC BY 4.0



אפס־ידע קלאסי שואל אם קיים סימולטור; "אפס־ידע אפקטיבי" שואל רק אם ספר הכללים שבחרתם יכול להוכיח ביעילות שסימולטור כזה אינו קיים. השאלה החלשה יותר היא שמאפשרת לבנייה לשמור על הודעה אחת, ללא הכנה מוקדמת ועם

נאותות מושלמת.

המבחן הישן: קיים סימולטור

הדרך הקלאסית לנסח אפס־ידע משתמשת בעוזר בדיוני שנקרא **סימולטור**.

הרעיון הוא כזה: דמיינו את ג'יין, שאינה יודעת **את הסוד של אליס**. אם ג'יין יכולה לייצר, לגמרי בעצמה, הוכחות שנראות בדיוק כמו ההוכחות שבוב היה מקבל מאליס, הרי שההוכחות של אליס לא לימדו את בוב שום דבר חדש. ג'יין כבר הייתה יכולה לזייף את אותה חוויה בלי הסוד של אליס.

לכן אפס־ידע קלאסי דורש סימולטור ממשי. צריך להיות אלגוריתם יעיל שיכול לייצר הוכחות מזויפות־למראה בלי לדעת את הסוד — את הער *(witness)*, בז'רגון; בסודוקו, העד הוא פשוט הלוח הפתור.

זו הגדרה חזקה, אבל בדיוק בנקודה הזאת נושכת אי־האפשרות הישנה. הנה האינטואיציה. הוכחה באמת לא־אינטראקטיבית היא בסך הכול מחרוזת. ברגע שלבוב יש את המחרוזת הזאת, הוא יכול להראות אותה למישהו אחר: הוא רכש את היכולת להוכיח את הטענה לאחרים, וכבר זה נשמע כמו יותר מ"שום דבר". המשפטים הקלאסיים מחדדים את האינטואיציה הזאת לכדי תוצאות אי־האפשרות שתוארו למעלה.

שלוש התכונות שעליהן המאמר מתעקש

אילוצים: שלושה מציינת המאמר כותרת

ללא אינטראקציה: אליס שולחת מחרוזת הוכחה אחת. אין פרוטוקול הלוך־ושוב.

ללא הכנה מוקדמת: אליס ובוב אינם מסתמכים על מחרוזת ייחוס משותפת ומהימנה או על אקראיות ציבורית אחרת שסוכמה מראש. מערכות רבות המכונות "אפס־ידע לא־אינטראקטיבי" עדיין מסתמכות על setup; במאמר הזה הכוונה היא לאפס הכנה מוקדמת.

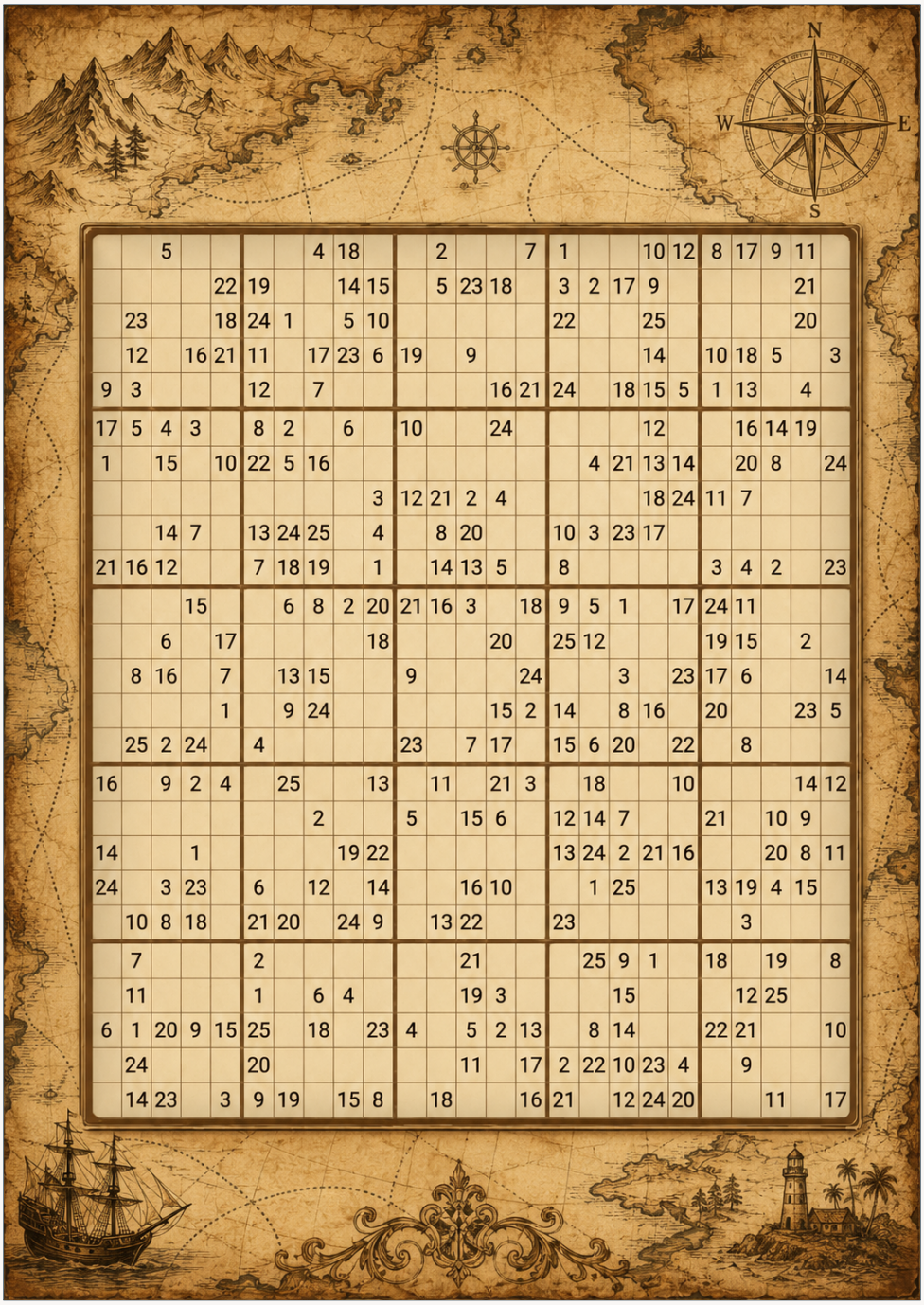
נאותות מושלמת: לטענה שקרית אין הוכחה תקפה. לא "כמעט אף פעם אינה מתקבלת"; אין שום הוכחה תקפה.

אלה בדיוק שלוש התכונות שיש למתמטיקה כתובה רגילה — וכפי שהוסבר למעלה, אפס־ידע קלאסי אינו יכול לשמור עליהן.

ההבדל דרך גרסת מגה־סודוקו

הנה דרך מפושטת בכוונה להרגיש את ההבדל.

אל תשתמשו בסודוקו רגיל של 9 על 9 לחלק הרציני של האנלוגיה. הוא קטן וסופי מדי: מחשב יכול פשוט לפתור אותו, או להוכיח שאין לו פתרון. במקום זאת דמיינו משפחה של חידות MegaSudoku(n). הגדילו את הכלל הרגיל: בחרו גודל בלוק n, הגדירו $N = n^2$, ובנו לוח N על N המחולק לבלוקים של n על n, עם N סמלים. סודוקו רגיל הוא רק המקרה הזעיר $n = 3$, $N = 9$: לוח 9 על 9, בלוקים של 3 על 3 ותשעה סמלים. סיפור מורכבות ההוכחות מתחיל רק כאשר מאפשרים ל-n לגדול, וכאשר הלוח יכול לכלול גאדג'טים נוספים שגורמים לו להתנהג כמו נוסחת SAT בתחפושת של סודוקו. נוסחת SAT היא פשוט רשימה של אילוצי כן/לא: האם אפשר להקצות למשתנים ערכי אמת/שקר כך שכל האילוצים יתקיימו?



סודוקו 25x25: אפשר לבדוק את כלליו בלי לחשוף את הלוח המלא — המחשה חזותית להוכחה שמאמתת פתרון מוסתר, כלומר את העד.

AI-generated editorial thumbnail — The Clean Paper CC BY 4.0

סודוקו ו-SAT: אותה חידה בשתי תחפושות

הכיוון ואת הכיוונים, בשני פועל התרגום מטאפורה. אינה "SAT נוסחת כמו "להתנהג יכול שסודוקו הטענה במלואו. לכתוב אפשר הקל

מסודוקו ל-SAT. SAT מדבר רק אמת/שקר, ולכן נותנים לו משתנה בוליאני אחד לכל שלשה (שורה, עמודה, ערך): *x(r,c,v)* פירושו "התא בשורה *r*, עמודה *c*, מכיל את הערך *v*." סודוקו 4 על 4 (בלוקים של 2 על 2, ערכים

1–4) זקוק ל-4·4·4 = 64 משתנים; סודוקו 9 על 9 הקלאסי זקוק ל-729. כל כלל של סודוקו הופך אז לקבוצה של פסוקיות. (פסוקית היא OR של משתנים או של שלילותיהם; הנוסחה כולה היא AND של כל הפסוקיות.)
בכל תא יש לפחות ערך אחד — פסוקית אחת לכל תא:

$$x(1,1,1) \vee x(1,1,2) \vee x(1,1,3) \vee x(1,1,4)$$

בכל תא יש לכל היותר ערך אחד — פסוקית "לא שניהם" לכל זוג ערכים:

$$\neg x(1,1,1) \vee \neg x(1,1,2) \quad \neg x(1,1,1) \vee \neg x(1,1,3) \quad \dots$$ וכך הלאה לכל ששת הזוגות.

בכל שורה מופיע כל ערך — עבור שורה 1 והערך 3: לפחות פעם אחת,

$$x(1,1,3) \vee x(1,2,3) \vee x(1,3,3) \vee x(1,4,3)$$

ולכל היותר פעם אחת: $\neg x(1,1,3) \vee \neg x(1,2,3)$, וכך הלאה לכל זוג תאים בשורה.
עמודות ובלוקים — קבוצות זהות; רק קבוצת התאים משתנה. עבור הבלוק השמאלי־עליון והערך 2:

$$x(1,1,2) \vee x(1,2,2) \vee x(2,1,2) \vee x(2,2,2)$$

ובנוסף פסוקיות "לא שניהם" לכל זוג.
הרמזים המוּדפסים — החלק הפשוט ביותר: כל רמז הוא פסוקית עם משתנה יחיד. הספרה 3 המודפסת בפינה השמאלית־עליונה הופכת לפסוקית

$$x(1,1,3)$$

ה-AND של כל אלה ניתן לסיפוק בדיוק כאשר לסודוקו יש פתרון — והשמה מספקת היא הפתרון: קוראים אילו $x(r,c,v)$ מקבלים אמת וממלאים את הלוח. עבור סודוקו 9 על 9 מתקבלים 729 משתנים וכמה אלפי פסוקיות, שפותר SAT מודרני מטפל בהם בתוך אלפיות שנייה. שימו לב לפסוקית הרמז $x(1,1,3)$: היא אומרת "התא הזה שווה בדיוק ל-3", ולא "כל התאים האלה שונים זה מזה" — אותה אסימטריה שתאלץ אותנו להשתמש בטריק נוסף עבור תאי הרמז בהערת הפרוטוקול בהמשך.
מ-SAT לסודוקו. המאמר זקוק לכיוון ההפוך והקשה יותר: בהינתן נוסחת SAT שרירותית, לבנות מגה־סודוקו שיש לו פתרון בדיוק כאשר הנוסחה ניתנת לסיפוק. הכללים הטבעיים של סודוקו יכולים לומר רק "התאים האלה שונים זה מזה", ולכן צריך לבנות אילוצים לוגיים שרירותיים — וזה בדיוק תפקידם של הגאדג'טים. גאדג'ט הוא אשכול קטן ומוכן מראש של תאים, אחד לכל פסוקית בנוסחה, שבו תאים ייעודיים ממלאים את תפקיד המשתנים (הסמל שהם מכילים מקודד אמת או שקר), והאילוצים הפנימיים של האשכול מתוכננים כך שהמילויים החוקיים היחידים שלו מתאימים להשמות שמספקות את אותה פסוקית. זו מלאכה סטנדרטית מהוכחות NP-completeness; עבור סודוקו מוכלל היא בוצעה בידי Yato ו-Seta בשנת 2003.
שני הכיוונים יחד אומרים שסודוקו N-על-N ו-SAT הם אותה בעיה בשתי תחפושות. זה מה שמאפשר למאמר הזה — ולמאמר המדעי — לספר סיפור על כל NP באמצעות לוחות וסמלים.

עדיין קל לדמיין את העד. אליס יודעת מילוי מלא וחוקי של המגה־סודוקו. בוב רוצה להשתכנע שקיים מילוי כזה, אבל אליס אינה רוצה לחשוף אותו. אם היא שולחת את כל המילוי, בוב משתכנע — אבל הסוד נעלם.

בגרסה הקלאסית של אפס־ידע, אליס ובוב מקיימים אינטראקציה. מודל מחשבתי ישן משתמש באריחים מכוסים. אליס מסתירה את הלוח הפתור, משנה בחשאי את שמות הסמלים לפני כל סבב, ונותנת לבוב לבדוק אילוץ מקומי אחד שנבחר באקראי: שורה, עמודה, תיבה או גאדג'ט. אם התאים שנפתחו מציגים סמלים שכולם שונים זה מזה, הביטחון של בוב עולה. אחר כך הכול מכוסה שוב ושמות הסמלים מוחלפים מחדש. (יש סיבוך אחד: הרמזים הנתונים של

החידה דורשים טריק נוסף, מפני ששינוי שמות הסמלים מסתיר גם אותם. ההערה שלמטה מסבירה כיצד הפרוטוקולים הקלאסיים פותרים זאת; התמונה הפשוטה מספיקה למה שיבוא בהמשך.)

איך הפרוטוקולים הקלאסיים באמת מטפלים בתאי הרמז

שונים האלה "התאים כולם אומרים והתיבות העמודות השורות, כללי עיוורת. נקודה יש השמות שינוי לטריק בדיוק מכיל הזה "התא אומר רמז אבל הסמלים. של שמות שינוי כל תחת נכון נשאר שונים כולם- מזה", זה אינו הוא .σ התמורה את לדעת בלי — כלשהו מוסווה סמל — $\sigma(5)$ רק רואה בוב השמות שינוי ואחרי 5", מוחלטת התעלמות תוך חוקי לוח איזשהו שקיים להוכיח יכולה אליס זאת, מתקנים לא אם דבר. לבדוק יכול סטנדרטיים. תיקונים שני יש הקלאסית בספרות .הזאת החידה על דבר מוכיח לא וזה המודפסים, מהרמזים **הפלטה.** מוסיפים ללוח המוסתר שורה נוספת של N תאים — פלטה שאליס ממלאת בסמלים 1…N בסדר ציבורי קבוע, ואז משנה את שמותיהם יחד עם כל השאר, כך שהיא מכילה $\sigma(1)…\sigma(N)$. כעת לאתגר האקראי של בוב יש אפשרות נוספת. מלבד בחירת שורה, עמודה, תיבה או גאדג'ט לפתיחה, הוא יכול לבחור **את הפלטה יחד עם תא רמז אחד**. אליס חושפת את שניהם; הפלטה חושפת את שינוי השמות של אותו סבב, ובוב בודק שתא הרמז מציג בדיוק את הגרסה ששמה שונה של הרמז המודפס. זה נשאר אפס־ידע מפני שבוב לומד רק את σ — שנבחרת מחדש באקראי בכל סבב וחסרת ערך בפני עצמה — ואת ערכו של תא שכבר הכיר מהחידה. שום דבר על התאים הסודיים אינו דולף, וסימולטור יכול לזייף את המראה פשוט באמצעות בחירת σ אקראית. הפרוטוקול בעל נאותות מפני שאליס מרמה נתפסת בהסתברות קבועה בכל סבב, וחוזרים על הסבבים עד שהספק נעשה זניח.

קימפול הרמזים החוצה. גרסה מבנית יותר מסירה את האתגר המיוחד במקום להוסיף אותו. במקום לאמת את ערך הרמז, כופים אותו באמצעות אילוצי אי־שוויון: מקשרים את תא הרמז לכל תא בפלטה מלבד התא שנושא את הערך שלו — "שונה מ-$\sigma(1)$, שונה מ-$\sigma(2)$, …, שונה מכל דבר פרט ל-$\sigma(5)$." הסמל היחיד שהתא יכול להחזיק באופן חוקי הוא הסמל של הרמז. כך כל אילוץ שוב הוא מהצורה "שני אלה שונים" — אינו משתנה תחת שינוי שמות, וניתן לבדיקה בדיוק כמו שורה. זה אותו מהלך שמשמש לקודקודים שצבועים מראש בפרוטוקול הקלאסי לצביעת גרפים, וזה גם רוח המילה גאדג'טים למעלה: בתמונת MegaSudoku-כ- SAT, הרמזים מקומפלים לגאדג'טים של אי־שוויון כמו כל אילוץ אחר.

הפרוטוקול הפיזי. פרוטוקול הקלפים הממשי לסודוקו (Gradwohl, Pinkas Naor, ו-Rothblum, 2007) אינו משתמש כלל בשינוי שמות, ומסדיר את הרמזים עוד לפני שההסתרה מתחילה. עבור כל תא, אליס מניחה שלושה קלפים זהים עם ערך התא — עם הפנים כלפי מטה עבור תאים סודיים, אבל **עם הפנים כלפי מעלה עבור תאי רמז**, כך שבוב רואה במו עיניו שהרמזים נשמרים לפני שהקלפים נהפכים. לאחר מכן קלף אחד מכל תא נכנס לחבילה של השורה שלו, אחד לחבילת העמודה שלו ואחד לחבילת התיבה שלו; כל חבילה מעורבבת ונחשפת, ובוב בודק שהיא מכילה את כל N הסמלים. הערבוב משמיד את מידע המיקום (זהו אפס־הידע), אבל הרמזים כבר נקבעו בזמן חלוקת הקלפים.

כך או כך, הלקח הוא אותו לקח שהמאמר הזה חוזר אליו שוב ושוב: פרוטוקול אפס־ידע הוא ניהול חשבונות קפדני של אילו עובדות שורדות את ההסתרה. שינוי שמות משמר "כולם שונים" ומוחק "שווה ל-5" — ולכן צריך להחזיר את "שווה ל-5" בדרך אחרת.

זה אינו הפרוטוקול שבמאמר. זהו המודל המחשבתי לאפס־ידע קלאסי:

- אליס ובוב מנהלים הלוך־ושוב.
- בוב בוחר בדיקות אקראיות.
- אליס חושפת רק עקביות מקומית, לא את הפתרון כולו.
- הוכחת הפרטיות פועלת בכך שמראים שאפשר היה לייצר את המראה שבוב רואה גם בלי הפתרון הסודי של אליס.

לכן אפס־ידע קלאסי בנוי סביב עובדה חיובית:

סימולטור באמת קיים.

כעת הסירו את החלקים הנוחים. אליס שולחת מחרוזת הוכחה אחת והולכת. אין הכנה מהימנה, אין מחרוזת אקראית משותפת שהוכנה מראש, ובוב אסור שיקבל אי פעם חידה שקרית. זו הסביבה שבה אפס־ידע קלאסי אינו יכול לשרוד.

לפני הטריק דרושה עוד דמות אחת. קבעו **ספר כללים**: מערכת הוכחה פורמלית, במובן של לוגיקנים — קבוצה קבועה של אקסיומות יחד עם כללים מכניים לבדיקת הוכחות מתמטיות כתובות. ZFC, מערכת האקסיומות הסטנדרטית של המתמטיקה, היא הדוגמה הקנונית. מכאן ואילך הכול נאמר ביחס לספר כללים שנבחר מראש, והבחירה גמישה: הבנייה פועלת עבור כל ספר כללים שתקבעו, כולל ZFC.

(הערת מינוח, בהשראת המאמר עצמו: "מערכת הוכחה" כאן פירושה תמיד ספר הכללים הזה — המערכת הפורמלית שבודקת הוכחות מתמטיות — ולעולם לא ההודעות שאליס שולחת. המנגנון של אליס ובוב נקרא "המוכיח והמאמת".)

הגרסה בסגנון גדל שומרת על סיפור המגה־סודוקו, אבל משנה את ההוכחה.

בחרו מערכת אילוצים שנייה באותו גודל מוצג, וקראו לה **D**. לצורך הסיפור, S ו-D הן שתי חידות MegaSudoku(n) באותו פורמט. מאחורי הקלעים, D עשויה להתחיל כנוסחה לוגית קשה בגודל אחר; אם צריך, אפשר לרפד אותה באילוצי דמה בלתי מזיקים כך שתתאים לאותו לוח. D בנויה מנוסחה לוגית שהיא למעשה **אינה ניתנת לסיפוק**: אין שום השמת ערכים שגורמת לכל אילוציה להיות נכונים, כשם שלחידה מקולקלת אין מילוי חוקי. דוגמת צעצוע תהיה נוסחה שדורשת גם X" אמת" וגם X" שקר". לכן ל-D אין מילוי תקף.

אבל D אינה יכולה להיות חידה מקולקלת ש-קל לחשוף. דוגמת הצעצוע שלמעלה נכשלת בכך: כל ספר כללים מפריך את X" וגם לא-X" בשורה אחת. D צריכה להיות שקרית באופן שספר הכללים שנבחר אינו יכול לאשר באמצעות טיעון קצר. אם ספר הכללים היה יכול להפריך את D בהוכחה קצרה, הסיפור שלמטה היה קורס: היה אפשר לשלול פורמלית את המסלול החלופי שאולי היה מאפשר לייצר הוכחות בלי הסוד של אליס, ואיתו גם את הבטחת הפרטיות. לכן בוחרים את D ממשפחה שספר הכללים הקבוע אינו יכול להפריך ביעילות: אין הוכחה קצרה, בתוך אותו ספר כללים, לכך של-D אין פתרון.

ההוכחה בהודעה אחת של אליס היא אז על טענת או/או:

או שלמגה־סודוקו האמיתי S יש פתרון, או שלפיתיון D יש פתרון.

זהו החיבור הלוגי. D **אינה** נוצרת בדרך קסומה שהופכת את S לנכונה. ההוכחה אינה טוענת "ל-D אין פתרון, ולכן ל-S יש פתרון." היא מוכיחה את הדיסיונקציה **S או D**. נאותות מושלמת אומרת שלדיסיונקציה שקרית לא יכולה להיות הוכחה תקפה. מכיוון ש-D שקרית במציאות — אין לה פתרון — הדרך היחידה שבה הדיסיונקציה יכולה להיות נכונה היא ש-S נכונה. לכן אם ההוכחה מתקבלת, ל-S חייב להיות פתרון. הפיתיון אינו יכול להפוך S שקרית לאמיתית.

אבל עבור החלק דמוי־אפס־ידע, שאלו מה היה קורה אילו ל-D כן היה פתרון. פתרון הפיתיון היה משמש כעד חלופי. הוא היה מאפשר למישהו לייצר הוכחות בלי לדעת את פתרון המגה־סודוקו האמיתי של אליס — כלומר, סימולטור. במציאות ל-D אין פתרון, ולכן מסלול הסימולטור הזה סגור. הנקודה היא שספר הכללים אינו יכול להוכיח ביעילות שהוא סגור.

לכן ל-D יש שני תפקידים. עבור **נאותות**, D שקרית, ולכן הוכחה תקפה של S" או "D מכריחה את .S עבור **אפס־ידע אפקטיבי**, קשה להפריך את ,D ולכן ספר הכללים אינו יכול לשלול במהירות את מסלול הפיתיון שהיה מאפשר סימולציה.

לכן מבחן האבטחה כבר אינו:

האם נוכל להוכיח שסימולטור באמת קיים?

הוא נעשה:

האם ספר הכללים שלכם יכול להוכיח ביעילות שהסימולטור אינו אפשרי?

אם התשובה היא לא, נובע משהו חזק באופן מפתיע: כל הבטחת אבטחה ש-(א) אפשר לצפות בה באמצעות הפעלת מבחן, ו-(ב) נובעת בהוכחה — בתוך אותו ספר כללים — מקיומו של סימולטור, מתקיימת בפועל. התקפה מוצלחת על אחת מהן הייתה בעצמה מסתכמת בהפרכה הקצרה החסרה, וההפרכה הקצרה הזאת אינה קיימת. זהו החלק ה"אפקטיבי" של אפס־ידע אפקטיבי.

לכן הניגוד בכיתה הוא:

אפס־ידע קלאסי: ההוכחות בטוחות מפני שסימולטור קיים.

אפס־ידע אפקטיבי בסגנון גדל: מתייחסים להוכחות כבטוחות עבור מבחני אבטחה נצפים מפני שספר הכללים אינו יכול להוכיח ביעילות שהסימולטור בלתי אפשרי.

הטענה השנייה חלשה יותר. זו גם הסיבה שהמאמר יכול לשמור על שלוש התכונות ששברו את הגרסה הקלאסית: הודעה אחת, ללא הכנה מוקדמת ונאותות מושלמת.

המבחן החדש: אי אפשר להוכיח שהסימולטור נעדר

ההרפיה של llango משנה את השאלה.

אפס־ידע קלאסי שואל:

האם קיים סימולטור?

אפס־ידע אפקטיבי שואל משהו חלש יותר:

האם ספר הכללים שבחרתם יכול להוכיח ביעילות שלא קיים סימולטור?

זה נשמע כמו התחמקות טכנית, אבל זהו הרעיון המרכזי. הבנייה חיה במצב מוזר: סימולטור למעשה אינו קיים — המאמר אומר זאת במפורש — אבל ספר הכללים שקבעתם אינו יכול להוכיח ביעילות שהוא אינו קיים. אם כל תוצאה רעה שמעניינת אתכם הייתה מחייבת הפרכה כזאת, המערכת עדיין מתנהגת כמו אפס־ידע ביחס לאותן תוצאות.

כאן נכנס גדל. לא כקישוט, ולא כ"גדל הופך קריפטוגרפיה לבטוחה". הקשר הוא תורת־הוכחתי. ספר כללים נקרא **אופטימלי** אם הוא, במובן מדויק, הטוב ביותר האפשרי: בכל פעם שספר כללים כלשהו יכול להפריך נוסחה מהסוג הרלוונטי בהוכחה קצרה, גם ספר הכללים האופטימלי יכול לעשות זאת, בהוכחה שאורכה גדול לכל היותר בגורם פולינומי. Krajíček ו-Pudlák שיערו בשנת 1989 ש-**לא קיימת מערכת הוכחה אופטימלית**: לא משנה איזה ספר כללים תקבעו, ספר כללים אחר יוכיח משפחה כלשהי של טענות אמיתיות באופן תמציתי בהרבה. זו אחת ההשערות הפתוחות המרכזיות במורכבות הוכחות, והיא בת־הדודה הסופית, במונחי תורת הסיבוכיות, של משפט אי־השלמות

של גדל: יש טענות אמיתיות שאין להן הוכחה קצרה בספר הכללים שקבעתם — לא מפני שאי אפשר להוכיח אותן עקרונית, אלא מפני שכל ספר כללים קבוע משאיר כמה טענות קצרות ואמיתיות בלי הוכחות קצרות.

המאמר מניח את ההשערה הזאת (בצורה מעט חזקה יותר של "**לעיתים אינסוף**" — often infinitely — המקובלת כאשר משתמשים בהשערות לצרכים קריפטוגרפיים). התמורה, לפי משפט של Krajíček ו-Pudlák, קונקרטית: לכל ספר כללים קיימת סדרה של נוסחאות שבאמת אינן ניתנות לסיפוק, אך ספר הכללים אינו יכול להפריך אותן בהוכחות קצרות — וחשוב מכך, אלגוריתם יעיל יכול **לייצר** אותן. התכונה האחרונה, אחידות (uniformity), היא שהופכת את כל הרעיון מטענת קיום לאלגוריתם ממשי שאליס יכולה להריץ: הפיתיונות D שלה יוצאים מפס ייצור, לא מן האוויר.

המהלך הקריפטוגרפי הוא להעמיד את המחסור הזה בכוח הוכחה לעבודה.

מה הבנייה עושה

הנה הבנייה של המאמר, כשהיא מופשטת לצורתה הבסיסית.

קבעו ספר כללים — למשל ZFC. תחת הנחת מורכבות ההוכחות, קיימת סדרה שאפשר לייצר ביעילות של נוסחאות שבפועל אינן ניתנות לסיפוק, אבל לספר הכללים אין הוכחה קצרה לכך שהן אינן ניתנות לסיפוק.

כעת בנו הוכחה בהודעה אחת מהצורה הזאת:

> או שהטענה האמיתית ניתנת לסיפוק, או שהנוסחה הקשה המיוחדת הזאת ניתנת לסיפוק.

הנוסחה הקשה המיוחדת אינה ניתנת לסיפוק. לכן אם מנגנון ההוכחה הבסיסי בעל נאותות מושלמת, קבלת ההודעה עדיין פירושה שהטענה האמיתית נכונה. כך מתקבלת נאותות מושלמת.

אבל עבור האבטחה דמוית־אפס־ידע, דמיינו שהנוסחה הקשה המיוחדת כן הייתה ניתנת לסיפוק. אז היה אפשר להשתמש בעד שלה כדי לסמלץ הוכחות בלי לדעת את העד האמיתי. הנוסחה אינה ניתנת לסיפוק במציאות — אבל ספר הכללים אינו יכול להוכיח זאת ביעילות. לכן הוא גם אינו יכול להוכיח ביעילות שהסימולטור בלתי אפשרי.

זהו הציר. המערכת אינה מסתירה את הסוד בכך שהיא מייצרת סימולטור קלאסי. היא מסתירה את הסוד, עבור מחלקה גדולה של מבחני אבטחה נצפים, מאחורי חוסר היכולת של ספר הכללים לאשר שהסימולטור נעדר.

מה המאמר טוען

המשפט המרכזי מגיע בכמה שכבות. תוצאת הליבה היא זו:

תחת הנחה קריפטוגרפית סטנדרטית — קיומן של **הוכחות לא־אינטראקטיביות שאינן מבחינות בין עדים** (non-interactive indistinguishable witness proofs), אובייקטים שנחקרו היטב ונובעים מכמה חבילות הנחות מבוססות — ותחת השערת מורכבות ההוכחות ש-**לא קיימת מערכת הוכחה אופטימלית** (**לעיתים אינסוף**), המאמר בונה, עבור כל בחירה של ספר כללים, מוכיח ומאמת בהודעה אחת עבור NP/SAT, עם **נאותות מושלמת** וללא הכנה מוקדמת, שהם אפס־ידע אפקטיבי ביחס לאותו ספר כללים. (NP/SAT הוא "המכנה המשותף הקשה ביותר" הסטנדרטי של בעיות דמויות חידה; מגה־סודוקו הוא אחת התחפושות שלו.)

עבור הטענה הרחבה יותר על שימור תכונות אבטחה ברות־הפרכה, המאמר מוסיף הנחה סטנדרטית נוספת, אמונת

הדה־רנדומיזציה BPP = P (בקירוב: אקראיות אינה מעניקה לאלגוריתמים כוח נוסף מהותי).

בתרגום משפת המשפטים:

- ההוכחה היא הודעה אחת.
- אין הכנה מהימנה.
- אי אפשר להוכיח טענות שקריות.
- המוכיח אינו אפס־ידע קלאסי — אין לו סימולטור.
- אבל כל תוצאת אבטחה ברת־הפרכה ומבוססת־משחק של אפס־ידע קלאסי ניתנת להשגה בסביבה הזאת.

"ברת־הפרכה" חשובה כאן. פירוש הדבר שכשל אבטחה ניתן לבדיקה באמצעות הרצת יריב בתוך משחק. הגדרות אבטחה קריפטוגרפיות רבות נראות כך: האם היריב יכול להבחין בין שתי הצפנות, להפוך פונקציה, לשחזר עד, או לנצח בניסוי מוגדר כלשהו? המשפט נותן מוכיח עבור כל תכונה ברת־הפרכה, אחת בכל פעם. מוכיח יחיד שנהנה מ-כל התכונות ברות־ההפרכה בעת ובעונה אחת כנראה בלתי אפשרי — מתקפת השימוש החוזר הישנה ("בוב יכול להראות את ההוכחה לאחרים") היא בעצמה תכונה ברת־הפרכה, והיא אכן נכשלת כאן. הצעת המאמר היא שמוכיח יחיד יכול, באופן סביר, לכסות את כל התכונות ברות־ההפרכה ה-טבעיות — אלה שמופיעות בפועל בפרקטיקה הקריפטוגרפית — אבל החלק הזה הוא משפט מותנה הנשען על מושג לא־פורמלי של "טבעי", יחד עם השערה מפורשת. ההבטחה מכוונת לכשלים נצפים, לא לכל משמעות פילוסופית או מבוססת־סימולציה של סודיות.

ראוי לציין מסקנה קונקרטית אחת: הבנייה מניבה את ההוכחות הלא־אינטראקטיביות הראשונות של הסתרת עד *(witness hiding)* עם מוכיח אחיד — "הוכחה שלחידה יש פתרון אינה עוזרת לך למצוא את הפתרון שלה", בלי אינטראקציה ובלי הכנה מוקדמת — אובייקט שנשמע צנוע, אך התחמק מבנייה במשך עשרות שנים.

מה זה לא אומר

זה החלק ששומר את הכתבה ישרה.

זה **לא** אומר שמשפטי אי־האפשרות הישנים היו שגויים. הבנייה עוקפת אותם באמצעות שינוי ההגדרה.

זה **לא** נותן אפס־ידע רגיל וקלאסי ללא אינטראקציה, ללא הכנה מוקדמת ועם נאותות מושלמת. המאמר אומר במפורש שלמוכיח שנבנה אין סימולטור.

זה **לא** אומר שאי אפשר להשתמש בהוכחה שוב. עדיין אפשר להראות הוכחה בת הודעה אחת למישהו אחר; המאמר אינו משמר תכונות בסגנון יכולת הכחשה .(deniability) (לאפס־ידע לא־אינטראקטיבי עם הכנה מהימנה יש אותה מגבלה.)

זה **לא** אומר שמדובר בפרוטוקול מעשי שמוכן לפריסה. זו תורת הסיבוכיות ויסודות הקריפטוגרפיה. התוצאה תלויה בהנחות מרכזיות ממורכבות הוכחות ומקריפטוגרפיה, והבנייה עוסקת במה שאפשרי עקרונית.

זה **לא** הופך את "גדל" לפרימיטיב אבטחה קסום. הקשר לגדל עובר דרך מערכות הוכחה, מערכות הוכחה אופטימליות ואנלוגים סופיים של אי־שלמות. האינטואיציה השימושית אינה "אי־שלמות מגינה על הסיסמה שלך". היא: אם ספר כללים אינו יכול להוכיח ביעילות שסימולטור בלתי אפשרי, אז אפשר לחסום, ברמת הגדרות האבטחה, מתקפות שהיו דורשות הוכחה כזאת.

למה זה בכל זאת מעניין

קריפטוגרפיה הופכת לעיתים קרובות קושי לבטיחות. פירוק לגורמים קשה, ולכן הנחות בסגנון RSA נעשות שימושיות. בעיות סריג קשות, ולכן קריפטוגרפיה מבוססת־סריג נעשית שימושית. כאן הקושי מוזר יותר: לא "קשה לחשב סוד", אלא "קשה להוכיח שאובייקט הוכחה מסוים אינו יכול להתקיים".

זו הסיבה שהמאמר מרגיש יוצא דופן. הוא מתייחס לאקסיומות ולספרי כללים כמעט כמו למשאבים קריפטוגרפיים. אי־האפשרות הרגילה אומרת שיש מתח בין נאותות לסימולציה. המהלך של llango מציב את המתח מאחורי וילון תורת־הוכחתי: הסימולטור נעדר, אבל המערכת הפורמלית אינה יכולה לחשוף את היעדרו ביעילות.

לקורא, החלק המפתיע אינו שהדבר יחליף את מערכות אפס־הידע של היום. כנראה שלא, לפחות לא ישירות. החלק המפתיע הוא שאפשר להשתמש באופן בונה במגבלה מן הלוגיקה המתמטית: לא רק כקיר, אלא כסוג של כיסוי.

עד כמה הראיות חזקות?

זהו מאמר של משפטים, ולכן "ראיות" פירושן כאן משהו אחר מאשר במאמר בביולוגיה או באסטרונומיה. השאלה אינה אם ניסוי שוחזר. השאלה היא אם ההגדרות, ההנחות ושרשרת ההוכחה תומכות בטענה.

ההוכחה פורמלית, והמאמר מפורש לגבי הנחותיו. אלה אינן הנחות אגביות. הוכחות לא־אינטראקטיביות שאינן מבחינות בין עדים הן אובייקטים סטנדרטיים בקריפטוגרפיה ונובעות מכמה חבילות הנחות מבוססות. השערת אי־קיומה של מערכת הוכחה אופטימלית היא השערה מרכזית במורכבות הוכחות. BPP = P היא אמונת דה־רנדומיזציה סטנדרטית, המשמשת רק למשפט הרחב יותר על תכונות ברות־הפרכה.

המאמר גם טוען שההנחות הן המחיר הנכון, לא פיגום שרירותי: הוא מוכיח כיוון הפוך שמראה שהן הכרחיות בעיקרן — אם בכלל קיימות בניות כאלה, אז חייבות להתקיים הוכחות לא־אינטראקטיביות שאינן מבחינות בין עדים, ובהינתן פונקציות חד־כיווניות סטנדרטיות, לא יכולה להתקיים מערכת הוכחה אופטימלית. וההנחות הן :"win-win" הפרכה של כל אחת מהן תהיה בעצמה תגלית פורצת דרך במורכבות הוכחות, בקריפטוגרפיה או בתורת הסיבוכיות.

אבל משום שהתוצאה מותנית, גם מידת הביטחון בה מותנית. אם ההנחות האלה ייכשלו, פירוש המשפט ישתנה. וגם אם ההנחות נכונות, ההבטחה אינה אפס־ידע קלאסי מלא; זו הגרסה המרוככת, תורת־הוכחתית, של המאמר.

לכן רמת הביטחון המתאימה היא גבוהה בכך שהמאמר מבסס תוצאת אפשרות מותנית וקוהרנטית; בינונית בכך שהנחותיו מתארות את העולם הקריפטוגרפי שבו אנחנו באמת חיים; ונמוכה לגבי כל השלכה מעשית מיידית.

למה זה חשוב

המאמר פותח נתיב שהיה אמור להיות סגור.

התיאוריה הקלאסית אומרת: אפס־ידע מלא אינו יכול להיות הודעה אחת ללא הכנה מוקדמת, ואינו יכול להיות בעל נאותות מושלמת. המאמר של llango אומר: אם נבקש את התוצאות של אפס־ידע שאפשר לבדוק במשחקי אבטחה, ואם נאפשר להגדרת האבטחה להיות תלויה במה שספר כללים יכול או אינו יכול להפריך ביעילות, אז אפשר לשחזר חלק גדול מן ההתנהגות השימושית — בהודעה אחת, ללא הכנה מוקדמת ועם נאותות מושלמת.

זה אינו תיקון קטן בהגדרה. זו דרך אחרת לחשוב על הבטחות קריפטוגרפיות. במקום לשאול רק מה קיים, שאלו מה ספר הכללים שלכם יכול לשלול. במקום להתייחס לאי־יכולת להוכיח כאל מטרד פילוסופי, השתמשו בה כמבנה.

העולם המעשי אולי לא ישתנה מחר. אבל המפה המושגית כן. יש כעת מובן פורמלי שבו "אף אחד אינו יכול להוכיח ביעילות שהסוד דלף" יכול להיות חזק מספיק כדי להשיב רבות מן ההגנות מבוססות־המשחק שרצינו מ"הסוד לא דלף".

לכן גדל שייך לכותרת.

בקצרה

הוכחות אפס־ידע מאפשרות למוכיח לשכנע מאמת שטענה נכונה בלי לחשוף את העד. תוצאות אי־אפשרות קלאסיות אומרות שאי אפשר לדחוס אפס־ידע להודעה אחת ללא הכנה מוקדמת, ושלא יכולה להיות לו נאותות מושלמת. המאמר של Ilango Rahul אינו מפריך את תוצאות אי־האפשרות האלה. הוא מגדיר מושג חלש יותר, אפס־ידע אפקטיבי: במקום לדרוש שסימולטור באמת יתקיים, הוא דורש שמערכת הוכחה שנבחרה — ספר כללים פורמלי כמו ZFC — לא תוכל להוכיח ביעילות שלא קיים סימולטור. תחת הנחות מרכזיות מקריפטוגרפיה (הוכחות לא־אינטראקטיביות שאינן מבחינות בין עדים) וממורכבות הוכחות (לא קיימת מערכת הוכחה אופטימלית), המאמר בונה מוכיחים בהודעה אחת עבור NP/SAT, ללא הכנה מוקדמת ועם נאותות מושלמת, שמשיגים את התוצאות ברות־ההפרכה ומבוססות־המשחק של אפס־ידע תכונה אחר תכונה. מוכיח יחיד שמכסה את כל התכונות ה"טבעיות" האלה הוא הרחבה נוספת, שחלקה השערתי — וכיסוי של ממש כל תכונה ברת־הפרכה כנראה בלתי אפשרי, מפני שהוכחות נשארות ניתנות לשימוש חוזר. התוצאה תיאורטית ומותנית, לא פרימיטיב שנפרס בפועל, אבל היא מראה דרך חדשה להשתמש באי־יכולת להוכיח, במובן של תורת ההוכחות, כמשאב קריפטוגרפי.

בדיקה מפוכחת

מה המאמר מראה: תחת ההנחות המוצהרות, אפשר לבנות מוכיחים בהודעה אחת, ללא הכנה מוקדמת ובעלי נאותות מושלמת עבור NP/SAT, שהם אפס־ידע אפקטיבי ביחס לכל מערכת הוכחה שנבחרה, ומשיגים כל תוצאה ברת־הפרכה ומבוססת־משחק של אפס־ידע קלאסי.

מה סביר אך לא הוכח ללא תנאי: שההנחות הנדרשות ממורכבות הוכחות ומקריפטוגרפיה נכונות. אלה הנחות רציניות ונחקרות היטב — והמאמר מראה שהן הכרחיות בעיקרן וגם מספיקות — אבל הן עדיין הנחות.

מה הוא אינו מראה: אפס־ידע קלאסי ללא אינטראקציה, ללא הכנה מוקדמת ועם נאותות מושלמת; מערכת מעשית שמוכנה לפריסה; יכולת הכחשה או אי־שימוש־חוזר בהוכחות; או שמשפט אי־השלמות של גדל, כשלעצמו, מאבטח קריפטוגרפיה.

המגבלות העיקריות: ההבטחה היא הרפיה של אפס־ידע; הגרסה הרחבה ביותר תלויה בכמה הנחות; הטענות על מוכיח אוניברסלי יחיד נשארות בחלקן השערתיות; והתוצאה היא בראש ובראשונה יסודית־תיאורטית.

כמה ביטחון צריך להיות לקורא כללי? גבוה בכך שמדובר בתוצאת תיאוריה מותנית חשובה, אם מקבלים את ההגדרות. בינוני בכך שההנחות אכן מתארות את המציאות. נמוך לגבי פריסה מעשית מיידית. המסקנה הבטוחה היא: המאמר אינו שובר את תוצאות אי־האפשרות של אפס־ידע; הוא מוצא דרך חדשה, תורת־הוכחתית, לעקוף את החלקים שלהן שחשובים למשחקי אבטחה רבים.

מקורות

2025/1296 ePrint IACR Ilango,
https://eprint.iacr.org/2025/1296

10.1109/FOCS63196.2025.00058 DOI ,2025 FOCS
https://doi.org/10.1109/FOCS63196.2025.00058