

## אוגדן בעיות מפורסמות במדעי המחשב

### על השימוש באלגוריתמים הסתברותיים לזיהוי מספרים ראשוניים

כתב גיל אבל

#### תיאור הבעיה

צייר יכול לבלות חיים שלמים בשיפור החיך של דמות שצייר. ביולוג יכול לחשוב ארוכות על דפוס הציור על כנפו של פרפר. משורר יכול להפוך ימים בשורה אחת בת שבע מילים שהעלה על הכתב. למתמטיקאים יש את המספרים הראשוניים הגדולים. למישהו מן החוץ אולי לא יובן העונג שבעיסוק במספרים הראשוניים, אך העוסקים בכך יטענו לסוג של שלמות שמספרים אלו נושאים בחובם, איכות אוניברסאלית נטולת זמן. אוניברסאליות שכזו, עד שסדרת מספרים ראשוניים נשלחה עם המסר שה "וויאגר" נשאה אל מחוץ למערכת השמש לציוויליזציות חוצניות. מדוע קיימים מספרים ראשוניים גדולים? כיצד הם מפורזים? מיהו המספר הראשוני מעבר לגדול ביותר המוכר כיום? מיהו הזוג הראשוניים בהפרש 2 זה מזה מעבר לגדולים ביותר המוכרים כיום?, ובעיית האב: כיצד מזהים מספר כראשוני?. מאז המאה השלישית לפנה"ס שאלות אלו ועוד שאלות רבות אחרות, ועל חלקן פרס כספי נכבד, גורמות לכל מתמטיקאי לחפוץ בגילוי כזה או אחר בעניין הראשוניים.

לראשוניות גם עניין מעשי מובהק. בין היתר, אלגוריתם ההצפנה בעל המפתח הגלוי RSA מסתמך לצורך קידוד מפתחותיו במספרים ראשוניים – גדולים ככול האפשר (ראה <http://cse.proj.ac.il/hebetim/june2003%5Crsa.htm>).

#### מספרים ראשוניים

תחילה נגדיר מספר שלם  $P$ , כך ש  $1 < P < D$  יקרא מספר ראשוני אם אין לו מחלק שלם  $D$  כך ש  $P > D > 1$ .

אין סוף מספרים ראשוניים : ניתן להוכיח כי קיימים אין סוף מספרים ראשוניים. אוקלידס (בערך 365 עד 275 לפני הספירה) בספרו "היסודות" הראה הוכחה על דרך השלילה: נניח ויש מספר סופי של מספרים ראשוניים. נכפול את כולם זה בזה ולתוצאה נוסיף 1. נקבל מספר  $X$  שאינו מתחלק באף אחד מהראשוניים בקבוצה ההתחלתית (כי נקבל שארית 1).  $X$  בוודאי שאינו מתחלק גם במספר פריק כלשהו, שהרי פריק שכזה היה מתחלק בעצמו לגורמיו הראשוניים, וגורמים אלו היו מחלקים את  $X$ . לכן  $X$  ראשוני, או שיש ראשוני נוסף שאינו בקבוצה ההתחלתית בו  $X$  מתחלק. בשני המקרים הגענו לסתירה, וההנחה נסתרת.

על מספר המספרים ראשוניים : מוערך כי בייצוג באורך של עד 512 ביטים קיימים כ  $10^{151}$  מספרים ראשוניים. לצורך השוואה : אם כל אטום ביקום היה מייצר ביליון מספרים ראשוניים כל מילי-שניה מאז היווצרות היקום ועד היום היו נוצרים רק  $10^{109}$  מספרים ראשוניים. מסד נתונים בו כל גרם מכיל 1 GB של מידע בו מאוכסנים כל המספרים הראשוניים הקיימים בייצוג באורך של עד 512 ביטים היה בעל מסה הגורמת ליצירת חור שחור – דבר שהיה מקשה על השימוש בו. העובדה כי קיימים מספרים ראשוניים כה רבים בשילוב העובדה כי קשה מאוד (כיום) לפרק מספר למרכיביו הראשוניים מאפשרת את שימושם בהצפנה.

## על אלגוריתמים הסתברותיים

בשפת הסימנים הסינית הציור למילה "משבר" הוא צירוף ציורי המילים "סכנה" ו"הזדמנות". לידתם של האלגוריתמים ההסתברותיים במשבר שכזה. בתחילת שנות ה 70 הראו רבין ופישר כי יש בעיות חישוביות אשר מספר הפעולות החישוביות הנדרש לפתרון הוא בלתי אפשרי לביצוע בזמן סביר. (בערך באותו זמן ובאופן עצמאי קוק, לוין וקרפ זיהו אוסף של בעיות קצת פחות קשות אך עדיין לכאורה בלתי פתירות באופן מעשי, np-complete). עננה של פסימיות וחוסר תוחלת העיבה על קהילת מדעי המחשב כאשר נראה כי ענפי מחקר שלמים אשר הכילו בעיות שנדמו כפשוטות אך מסתבר כי היו קשות מדי לפתרון עומדים להכרת. אולם בהרצאה בסטוקהולם ב 1974 העלה רבין, באותה הרצאה בה דבר על הבעיה, כלי אשר אפשר להתמודד לפחות עם חלק מהשלכותיה :

"I proposed we should give up the attempt to derive results and answers with complete certainty. We should use randomness in a certain way and get the results more quickly but with a small probability of error ..." (M. O. Rabin)

"הצעתי כי עלינו לוותר על הניסיון להגיע לתוצאות ותשובות וודאיות. עלינו להשתמש באקראיות באופן מסוים ולהגיע לתוצאות יותר מהר אך עם סיכוי קטן לשגיאה." (מיכאל א. רבין)

שימוש באקראיות היה קיים כבר במדעי המחשב אולם לא בצורה שכוז, אלא לצורך יצירת סימולציות להתנהגות סטוכסטית בלבד (התנהגות סטוכסטית : צורת התנהגות בעלת מרכיב אקראי בה ידיעת העבר וההווה אינה מספיקה לחיזוי העתיד, למשל תנועת בראון או התנהגות הבורסה).

הרעיון לוותר על הנכונות המוחלטת, לבנות אלגוריתם הסתברותי שאינו דטרמיניסטי, אשר במהלך עבודתו "מטיל מטבעות" ופועל עפ"י התוצאות, אלגוריתם אשר בהסתברות גבוהה יתן תוצאה נכונה, אולם עלול גם לטעות, היה חדשני לחלוטין.

ניתן להדגים את הרעיון באופן הבא : בהינתן מיכל שוקולד גדול, אנו רוצים לבדוק נוכחות עובש. מניסויים קודמים במפעל ידוע כי בלקיחת דגימה של 1 גרם יש סיכוי של 80% לבחור דגימה המכילה נבגי עובש אשר יזוהו. בהתאם, תהליך זיהוי הכולל לקיחת 5 דגימות אקראיות יתן תשובת false-negative (תשובה שלילית לנוכחות זיהום עובש למרות שהמיכל מזוהם) בהסתברות של  $0.2^5 = 0.00032$ . זהו אלגוריתם הסתברותי לזיהוי עובש אשר עלול לטעות וניתן לחשב את ההסתברות לטעות. הטלות המטבע כאן היא הבחירה האקראית בדגימה, ומאותו מיכל, בהפעלה נוספת של אלגוריתם הזיהוי, תתכן תוצאה שונה !

רבין אמר ועשה – ב 1975 ב M.I.T , בהסתמך על תוצאות של מילר, יצר רבין את האלגוריתם המהיר ביותר לזיהוי מספרים ראשוניים שנוצר עד אז, ועד היום. כבר בערב הראשון להפעלתו זוהו מספרים ראשוניים וזוגות ראשוניים בהפרש 2 גדולים מכול אלו שזוהו עד אז. אלגוריתם זה הינו האלגוריתם העיקרי אשר בשימוש כיום לצורך זיהוי מספרים ראשוניים, והינו אלגוריתם הסתברותי.

כיום כשליש מסך האלגוריתמים המתפרסמים הם הסתברותיים , והאלגוריתמים הסתברותיים נפוצים ברובוטיקה, חישוב מבוזר, תקשורת, איחזור מידע ממאגרים, הצפנה ואפילו בענף הפריצה למערכות מחשב.

## מבחן רבין-מילר לראשוניותו של מספר

מבחן רבין-מילר לראשוניותו של מספר מתבסס על קיום עדים  $A$  אשר בקיימם מערכת יחסים מסוימת עם המספר הנבדק (סעיף 3) קובעים אותו כראשוני. מערכת יחסים זו מסתמכת על היפוטיזת ריימן ותורת המספרים.

המבחן הוא הסתברותי. רבין ומונייר הראו בשנת 1980 כי מספר פריק יעבור את המבחן לכול היותר ב 25% מה  $A$  האפשריים (במקרה הגרוע ביותר. במעשי מדובר על פחות מאחוז). ראשוני יעבור לכל  $A$  אפשרי. מכאן כי ביצוע  $i$  חזרות יביא לטעות בהסתברות  $0.25^i$  לכל היותר. בביצוע 1000 חזרות נקבל הסתברות לטעות הקטנה בהרבה מההסתברות שתגרם טעות עקב כשל חומרתי מיקרי במחשב !

בהנתן מספר  $N$  אותו תרצה לבדוק לראשוניות :

דוגמא : נבדוק את  $N=13$ .

1. חשב את ה  $R$  המקסימלי ואת  $S$  כך ש :  $N = 1 + 2^R * S$ ,  $S$  שלם ואי-זוגי.

נחשב :  $R=2, S=3$

2. בחר  $A$  כלשהוא,  $1 < A < N - 1$ ,

נבחר  $A=5$

3. בדוק עדות : אם

תנאי ראשון:  $A^S \bmod N = 1$

התנאי הראשון אינו מתקיים :  $5^3 \bmod 13 = 125 \bmod 13 = 8$

או

תנאי שני :  $A^{2^j * S} \bmod N = -1$  בעבור  $J$  כלשהו בין 0 ל  $R-1$  (כולל)

התנאי השני מתקיים !

בעבור  $J=1$ ,  $5^{2*3} \bmod 13 = 5^6 \bmod 13 = 15,625 \bmod 13 = -1$

אזי  $N$  מוערך כראשוני !

ביצוע המבחן ב PC רגיל למספר בן 100 ספרות עורך שבריר שניה, כפי שניתן להתנסות באתרים שונים ברשת.

(למשל : <http://www.cryptomathic.com/labs/rabinprimalitytest.html>)

בי"עולם האמיתי" יצירת מספר ראשוני הינה כדלקמן :

1. ייצר מספר בן  $N$  ביטים (כ  $N$  המבוקש).

2. קבע את ה high order bit וה low order bit להיות 1 (בכדי לוודא כי המספר בגודל המתאים ואי-זוגי, בהתאמה).

3. בדוק ב "brute-force" כי המספר אינו מתחלק בראשוניים קטנים (נהוג ראשוניים עד 256) – צעד זה חושף 76% מהמספרים הפריקים.
4. ביצוע מבחני רבין-מילר ברמת הבטחון הנדרשת. אם המספר עבר את כל המבחנים – השתמש בו כראשוני.

מציאת מספר ראשוני  $N$  גדול נחשבה במשך זמן לבעיה שפתרונה דורש לפחות  $O(N^{1/2})$  פעולות חילוק. כיום קיימים מספר אלגוריתמים לבדיקת ראשוניות, בתנאים מסוימים חלקם אף דטרמיניסטיים (קיימת גם הרחבה מותנית דטרמיניסטית למבחן רבין-מילר), אך מבחן רבין-מילר ההסתברותי הוא עדיין הנפוץ והשמיש ביותר.

### מיכאל א. רבין

מיכאל רבין נולד בשנת 1931 בברסלאו שבגרמניה. בגיל 4 עלה עם משפחתו ארצה והשתקע בחיפה. למרות היותו מבית דתי התעקש רבין ושכנע את אביו לשולחו לבית הספר הריאלי בחיפה. בתחילת שנות ה 50 החל רבין בלימודיו באוניברסיטה העברית – בדיוק בזמן שהמאמרים הראשונים על מחשבים פורסמו בישראל. שבוי בקסמי כתבי טורינג וגדל ובחוסר מחשבים ותשתית בארץ פנה רבין לארה"ב וסיים דוקטורט בפרינסטון, כתלמידו של צ'רץ. בקיץ 1957 בחופשת לימודים ב IBM, הגו רבין וחברו דנה סקוט את רעיון האוטומט הלא דטרמיניסטי, ותוך מספר שבועות פרסמו מאמר בעניין, מאמר שהפך להיות אחד מהמצוטטים ביותר בתחום מדעי המחשב. שנה לאחר מכן החל בעבודתו על "קושי של חישובים". בשנות השבעים עסק רבין בנושא האי-חישוביות מחד ובנושא החישוב ההסתברותי מאידך ויצר את אלגוריתם רבין-מילר. תוך כדי כל זאת הקים ב 1958 את המחלקה למדעי המחשב באוניברסיטה העברית ושירת כרקטור האוניברסיטה העברית. כיום רבין מחלק זמנו בין אוניברסיטת הארוורד לאוניברסיטה העברית, ועוסק באלגוריתמים הסתברותיים להבטחת אמינות במחשבים מקביליים גדולים. רבין הינו בין היתר חתן פרסי ויצמן למדע, רוטשילד למתמטיקה וחתן פרס טורינג (משנת 1976).

### מקורות

Denis Shasha and Cathy Lazere. Out of Their Minds: The Lives and Discoveries of 15 Great Computer Scientists, NY: Copernicus, 1995, pp. 68-88.