עוד כמה מילים על מספרים ראשוניים

המספרים הראשוניים עוררו את סקרנותם של בני האדם עוד משחר ההיסטוריה המדעית. התבוננות באופי הכאוטי של פיזורם מעוררת תמיהות רבות ואכן כמה מהבעיות הפתוחות הידועות ביותר במתימטיקה עוסקות במספרים הראשוניים.

כבר היוונים ידעו שיש אינסוף מספרים ראשוניים (הוכחה: אם מניחים ש $p_1,\dots,p_n$ היא הרשימה המלאה של כל המספרים הראשוניים אז מגיעים לסתירה מתוך התבוננות במספר $\Pi p_i + 1$. השלימו את הפרטים). האם ניתן לומר דבר מה על הפיזור שלהם? זו שאלה מרתקת המובילה לכמה מההישגים היפים ביותר של המתמטיקה וגם לכמה מהבעיות הפתוחות המפורסמות ביותר בה. למשל: עד כמה צפופים יכולים המספרים הראשוניים להיות. בפרט, בין המספרים הראשוניים הקטנים לא קשה למצוא זוגות של ראשוניים שההפרש ביניהם הוא 2 (אלה קרויים תאומים) כגון 17 ו- 19 או 59 ו- 61. זוהי בעייה פתוחה אם יש אינסוף זוגות של תאומי ראשוניים.

מאידך גיסא, לא קשה למצוא מספרים גדולים כרצוננו N כך שבין N ל- $N+\Omega\left(\frac{logN}{\log\log N}\right)$ אין אף מספר ראשוני ("מדבריות" של ראשוניים). למשל: בוחרים k כלשהו ומבחינים שכל המספרים k!+2, k!+3, ..,k!+k פריקים (השלימו את הפרטים).

בכיוון השני ניתן לשאול עד כמה יכולים המדבריות הנ"ל להתארך. בעייה פתוחה מפורסמת במיוחד (אולי הידועה ביותר בכל המתימטיקה) היא השערת רימן שבאחד מניסוחיה אומרת שלכל $\varepsilon>0$ אם x גדול מספיק אז יש מספר ראשוני בקטע שבין x ל- $x+x^{1/2+\varepsilon}$. תוצאה חיובית בכיוון זה שניתן להוכיח בשיטות אלמנטריות היא שלכל n יש מספר ראשוני בין n ל-$2n$. בצד ההישגים ראוי לציין את משפט המספרים הראשוניים שהוכח בתחילת המאה ה- 20. הוא מסביר מהי הצפיפות הממוצעת של המספרים הראשוניים. תצפית קלה במספרים הראשוניים תראה לכם שצפיפותם הולכת ויורדת ככל שהמספרים הולכים וגדלים, אולם נשאלת השאלה באיזה קצב. אם נסמן ב- $\Pi(N)$ את מספר המספרים הראשוניים הקטנים מ- N אז המשפט הנ"ל אומר כי $\Pi(N)\sim\frac{N}{\ln N}$ כלומר, $\lim_{N\to\infty}\frac{\Pi(N)\cdot\ln N}{N}=1$. יוצא למשל שאם אנו בוחרים באקראי במספרים בעלי 700 ספרות בינריות, ניתן לצפות שבערך $\frac{1}{500}$ מהם יהיו ראשוניים. נניח עתה לרקע המתימטי ונחזור לדיון האלגוריתמי.

מבחני ראשוניות

מרכיב חיוני באלגוריתם RSA ובאלגוריתמים מודרניים נוספים הוא מבחן הראשוניות. מדובר בבעייה היסודית הבאה:

אנו מקבלים כקלט מספר טבעי n ועלינו להכריע אם הוא ראשוני או פריק. אנו נציג אלגוריתם הסתברותי (של Miller ורבין) לפתרון הבעייה. מעניין להעיר שבתחום זה פותח לראשונה הרעיון של אלגוריתם הסתברותי (ע"י רבין ו- Solovay Strassen) מה שהפך בהמשך לאבן פינה בתורת האלגוריתמים. יש לציין שרק לאחרונה (ב- 2003) נמצא האלגוריתם הפולינומי הדטרמיניסטי הראשון לבחינת ראשוניות (Agarural, Saxena, Kayal).

מבחן הראשוניות **Rabin-Miller**

טענה 1: אם יש $1<a<n$ כך ש $a^{n-1}\neq 1 \pmod{n}$ אז n אינו ראשוני. זהו כמובן ניסוח חילופי למשפט פרמה.

טענה 2: אם למשוואה $x^2\equiv 1 \pmod{n}$ יש שורשים מלבד $x=\pm1$ אז n אינו ראשוני.

הוכחה: מעל שדה, יש לפולינום ממעלה d לכל היותר d שורשים.

האלגוריתם: בוחרים a באקראי. אם $gcd(a,n)>1$ עונים "n פריק". מסמנים $n-1=2^t\cdot k$. מחשבים את המספרים $a^k,a^{2k},a^{4k},a^{8k},\dots,a^{n-1} \pmod{n}$. אם $a^{n-1}\neq 1 \pmod{n}$ עונים "n פריק". אם $a^{n-1}\equiv 1 \pmod{n}$ מחפשים את $a^{2^j k}$ האחרון בסידרה השונה מ- 1. אם זה איננו 1- שוב n אינו ראשוני ועונים "n פריק".

בכל מקרה אחר עונים "n כנראה ראשוני"

זהו האלגוריתם ההסתברותי הראשון שאנו מנתחים בקורס וראוי להסביר כמה מהעקרונות המשתקפים כאן ומופיעים בניתוח של אלגוריתמים הסתברותיים רבים.

לאלגוריתם הזה יש "שגיאה חד צדדית". כלומר כאשר האלגוריתם עונה "n פריק" הוא צודק בוודאות. כאשר התשובה היא "n כנראה ראשוני" זוהי טענה הסתברותית בלבד. נדייק: אם n פריק, ההסתברות (ע"פ הבחירות של a) שנגיע למסקנה ש-n כנראה ראשוני היא $\geq 1/2$.

במקום זה ראוי להסביר כמה נקודות עקרוניות:

א. a כזה שבשבילו האלגוריתם מסיים את הריצה בהודעה "n פריק" נקרא עד (witness) לפריקות של n. הרעיון הבסיסי ביותר שביסוד האלגוריתם הוא שניתן להוכיח (כפי שנעשה להלן) שיש עדים רבים מאוד ואם n פריק אז לפחות מחצית מהמספרים $1 < a < n$ הם עדים לפריקות של n. אומנם איננו יודעים כיצד למצוא עד כזה באופן שיטתי וודאי אבל אנו מסתמכים על ריבויים של העדים ודוגמים a באקראי בידיעה שאם n פריק אז הסיכוי לגלות זאת בעזרת a (כלומר ההסתברות לכך ש- a עד לפריקות של n) הוא לפחות 1/2.

ב. יש מקום לתהות אם הסתברות הצלחה של 1/2 מספקת. ואכן, כרגיל לא נסתפק בכך. אם קיבלנו את התשובה "n פריק" ידוע לנו שהתשובה נכונה ואנו עוצרים את ריצת האלגוריתם. מאידך גיסא, אם התשובה היא "n כנראה ראשוני" רצוי לחזור ולדגום a אחר. נניח שנחזור על כך נאמר 20 פעם ובכל פעם התשובה היא "n כנראה ראשוני". ההסתברות שרק מזלנו הרע פועל נגדנו (והאמת היא ש- n פריק) היא קטנה מ- $(1/2)^{20} < 10^{-6}$. מובן שניתן להקטין במחיר חישובי נמוך מאוד את ההסתברות לסוג כזה של טעות כאוות נפשנו.

ג. בדיון רחב יותר באלגוריתמים הסתברותיים (שאיננו עושים בקורס הנוכחי) מתעוררות אפשרויות מעניינות אחרות. כפי שראינו, באלגוריתם רבין- Miller יש אפשרות לטעות חד צדדית ואנו יכולים להקטינה כרצוננו. באלגוריתמים הסתברותי-ים אחרים (לבעיות במגוון רחב של תחומים) נתקלים לעיתים גם בטעויות דו צדדיות. אולם, יש גם תופעה אחרת. במקרה שלנו זמן הריצה של האלגוריתם כמעט ואינו מושפע מהבחירה האקראית שעשינו (של המספר a) והוא תלוי תלות חלשה (לוגר-יתמית) בדרגת הוודאות הסטטיסטית הרצויה לנו (דרך מספר החזרות שנעשה על האלגוריתם). אולם, יש גם אלגוריתמים הסתברותיים שבהם זמן הריצה תלוי בבחיר-ות האקראיות שלנו ("זמן הריצה הוא משתנה מקרי"). המצבים הרצויים לנו הם כאלה שבהם תוחלת זמן הריצה היא קטנה. מוטב מזה שההסתברות לזמן ריצה ארוך היא אפסית. (וכמובן טוב מזה המצב כמו כאן שבו זמן הריצה סביר תמיד).

כדי לסכם זאת נאמר שהאקראיות יכולה להשפיע בשני אופנים עיקריים:

א. נכונות התשובה (נכונה תמיד, נכונה בהסתברות גבוהה, סוג אחד של תשובה נכון תמיד, סוג אחר רק בהסתברות גבוהה וכו').

ב. זמן הריצה (קצר תמיד, קצר בממוצע, קצר כמעט בוודאות וכו').

הערה חשובה נוספת היא ש"כרגיל" מבחן הראשוניות יכול להיעשות בהצלחה כבר על סמך הקריטריון הראשון (טענה 1 לעיל). לרוב המספרים הטבעיים הפריקים n, לפ-חות מחצית מן המספרים $1 < a < n$ הם עדים לפריקות של n במובן זה שהם מראים כי משפט פרמה מופר. מאידך גיסא, יש מספרים שלמים פריקים (שהקטן מביניהם הוא $561 = 3 \cdot 11 \cdot 17$) הנקראים "מספרי Carmichael". שאף אחד מן המספרים $1 < a < n$ הזרים ל- n אינו עד לפריקותם במובן זה. זוהי הסיבה שנדרשת גם הטענה השנייה. אנו נוכיח טענה מוחלשת מהסוג הדרוש לנו:

טענה: אם n אינו מספר קרמייקל אז לפחות מחצית מן המספרים $1 < a < n$ הם עדים לפריקות של n.

הוכחה: נתחיל בצעד הראשון של האלגוריתם. ברור שכל $1 < a < n$ שאינו זר ל- n (כלומר gcd(a,n)>1) הוא עד לפריקות של n. אנו נראה שגם מבין המספרים $1 < a < n$ הזרים ל- n, לפחות מחצית הם עדים לפריקות של n. ואכן, נתבונן בקבוצה $S = \{1 < a < n \mid gcd\,(a, n) = 1, a^{n-1} \equiv 1 (mod n)\}$.

S חלקית לחבורה הכפלית מודולו n. יתר על כן S אינה כל החבורה מפני שע"פ ההנחה n אינו מספר קרמייקל. זאת ועוד: S היא חבורה, מפני שכפי שקל לוודא, היא

סגורה לכפל ולהופכי. כזכור עוצמתה של תת חבורה מחלק את עוצמת החבורה, ובפרט גודלה לכל היותר מחצית מגודל החבורה כולה.

אבל S כוללת בדיוק את המספרים שאינם עדים לפריקות של n, וכפי שראינו $|S| \leq \frac{\phi(n)}{2} \leq \frac{n-1}{2}$, והטענה הוכחה.