Butlletí de la Societat Catalana de Matemàtiques
Vol. 17, núm. 2, 2002. Pàg. 21–28

L'algorisme de primalitat AKS*

Josep Díaz

En aquest article vull presentar una perspectiva històrica de la recent i sorprenent descoberta de Manindra Agrawal, Neeraj Kayal i Nitin Saxena d'un algorisme polinòmic per al test de primalitat [4]. Aquest article també es pot trobar a `http://www.cse.iitk.ac.in/users/manindra/index.html`

El resultat va merèixer una columna al *New York Times* el proppassat 8 d'agost del 2002.

Cal remarcar que l'existència d'algorismes polinòmics per al problema de la primalitat no implica que s'hagi trobat una manera eficient de factoritzar un enter compost, de manera que l'algorisme AKS (Agarwal-Kayal-Saxena) no afecta la seguretat dels sistemes criptogràfics RSA (`http://world.std.com/~franl/crypto/rsa-guts.html`) o PGP (`http://www.pgp.com/`).

1 Els fonaments

Al començament dels setanta, S. Cook [8] i R. Karp [12] van definir les classes de complexitat P, NP i NP-completa. Informalment, un problema és a la classe P si pot ser resolt amb un algorisme determinista en temps (nombre de passos) polinòmic respecte a la grandària de l'entrada. Per exemple, utilitzant l'algorisme que s'ensenya a l'escola, el temps per multiplicar dues matrius d'enters $n \times n$ és $O(n^3)$. (Recordeu que, donades dues funcions f i g, la notació $f(n) = O(g(n))$ significa $\exists c > 0 \;\; \exists n_0 \in \mathbb{N} \;:\; \forall n \geq n_0 \;\; f(n) \leq cg(n)$.) Per introduir la classe NP necessitem el concepte d'*indeterminisme*. Donat un problema i una entrada de grandària n, un algorisme indeterminista *conjectura* una solució, i verifica si la solució conjecturada és realment una solució vàlida per al problema. La classe NP és la classe de problemes que poden ser resolts per un algorisme indeterminista en temps polinòmic (en la grandària

* Una versió preliminar d'aquesta exposició va aparèixer al *Bulletin of the EATCS, 78* (octubre 2002).

de l'entrada). Aquí el temps és el nombre de passos que triga a comprovar si la solució conjecturada és realment una solució, sense atorgar cap cost a la generació de la conjectura.

Per exemple, considerem el problema COMPOST, que consisteix a decidir si, donat un enter N, hi ha dos enters $r > 1$ i $s > 1$ tals que $N = rs$. Un algorisme indeterminista per a COMPOST «conjectura» r i comprova si $r|N$ (r divideix N), la qual cosa es pot fer en $O(\log^2 N)$ operacions. És important fer notar que, quan tractem amb problemes on l'entrada és un enter N, la grandària n és $\log N$ (el nombre de bits necessaris per a escriure N). Per tant, qualsevol algorisme que hagi de considerar els $N - 1$ enters previs, necessita com a mínim un temps $O(N) = O(2^{\log N})$, es a dir, exponencial en la grandària n!

La classe NP-Completa és la classe dels problemes més difícils de resoldre en temps polinòmic a la classe NP, en el sentit que si es demostra que un problema a NP-Completa té solució polinòmica determinista això implicaria que tots els problemes NP serien també a P, i per tant $P = NP$. La verificació d'aquesta igualtat o de la seva negació és el problema del milió de dòlars (vegeu, per exemple, `http://www.claymath.org/prizeproblems/pvsnp.html`). Aquestes classes de complexitat, igual que les classes de complexitat superior, l'anomenada *jerarquia polinòmica*, es defineixen codificant els problemes com a llenguatges formals i utilitzant màquines de Turing. El lector interessat pot consultar qualsevol dels nombrosos textos sobre el tema [14, 10].

El problema PRIMERS consisteix a *decidir* si un enter donat N és primer. La demostració que PRIMERS pertany a NP no és trivial del tot i es deu a V. Pratt [15].

Des de l'antiguitat, molta gent ha treballat en el problema de decidir si un enter és primer o no i, en cas que no, trobar la seva factorització en nombres primers. És sabut que al segle X aC els xinesos creien tenir un test per decidir el problema PRIMERS: N és primer si i només si divideix $2^N - 2$. El test falla ja per a $N = 341$. Al segle XVII, Fermat va enunciar una generalització de la conjectura xinesa en una direcció. El resultat es coneix com a Petit Teorema de Fermat: si un enter N és primer, aleshores per a qualsevol $a \neq 0$ se satisfà

$$a^{N-1} \equiv 1 \pmod{N}. \tag{1}$$

Si es disposés de la direcció oposada del petit teorema de Fermat, un test de no-primalitat d'un enter donat N consisitiria a verificar l'existència de $a \in \{1, \ldots, N-1\}$ tal que $a^{N-1} \not\equiv 1 \pmod{N}$. Hi ha enters compostos, però, que satisfan l'equació (1), per exemple $N = 561$. Aquests enters estranys s'anomenen *nombres de Carmichael*.

Per tal de dissenyar un algorisme indeterminista per reconèixer PRIMERS, Pratt dissenyà un sistema lògic amb dues regles d'inferència: l'una basada en el Petit Teorema de Fermat i l'altra basada en el fet que, si $N > 3$ és primer, aleshores $N - 1$ és compost, i si $N - 1 = p_1 p_2 \cdots p_k$ és la seva descomposició en factors primers, per a qualsevol a, $1 < a < N$, $x^{(N-1)/p_j} \not\equiv 1 \pmod{N}$ per a algun p_j. Finalitza demostrant que en un nombre de passos $O(\log^4 N)$, i

utilitzant un sistema indeterminista, pot derivar un teorema decidint si N és primer, i per tant demostrar que PRIMERS $\in NP$.

Notem que COMPOST és el problema complementari de PRIMERS, cosa que implica que PRIMERS pertany a $NP \cap co-NP$, on la classe $co-NP$ és la classe de problemes tals que el seu complementari està a NP. Aquest fet va fer que V. Pratt conclogués el seu article manifestant:

> *We advocated membership in $NP \cap co-NP$ as a strong reason for presuming non-NP-completeness of* PRIMES...

ja que si PRIMERS fos NP-complet, aleshores $NP = co-NP$ contràriament a la conjectura establerta que $NP \neq co-NP$.

L'any 1979, Garey i Johnson [10] van presentar dos problemes que es coneixia que eren a $NP \cap co-NP$: *Programació Lineal* i PRIMERS. Actualment, ja ha estat demostrat que els dos pertanyen a la classe P. Això sembla invitar a fer una conjectura del tipus $P = NP \cap co-NP$. Considerem, però, el problema π_1: *donat com a entrada un parell d'enters (n,r) decidir si existeix un $s < r$ tal que $s|n$*. Aquest problema és a NP: conjecturem s i verifiquem que $s < r$ i que $s|n$. El complementari $\bar{\pi}_1$ de π_1 és: *donat com a entrada un parell d'enters (n,r), decidir que tots els factors primers no-trivials de n són més grans que r*. Aquest problema també és a NP: conjecturem la factorització primera única de $n = p_1, \ldots, p_k$ i verifiquem que tots els factors són primers més grans que r. Per tant, $\pi_1 \in NP \cap co-NP$, però és fàcil veure que π_1 correspon al problema de *factoritzar n*, que pertany a $NP \cap co-NP$ i no es coneix si el problema és NP-complet o és a P. Personalment, tinc els meus seriosos dubtes que el problema de la factorització d'un enter estigui a P. En tot cas, un algorisme determinista polinòmic per a π_1 sí que posaria en perill els sistemes criptogràfics RSA i PGP.

2 El passat recent

Els anys següents de l'aparició de l'algorisme indeterminista de Pratt, els esforços per trobar un algorisme determinista pel problema PRIMERS bàsicament intenten determinitzar el procediment de Pratt. L'estratègia per decidir la primalitat d'un enter N consistia a cercar determinísticament el *testimoni a* (que Pratt conjecturava indeterminísticament) del fet que N és compost. Si no es troba cap testimoni, aleshores N és primer. Com a criteri per a testimonis, E. Miller [13] i M. O. Rabin [17] feien servir l'equació (1) en combinació amb el següent resultat de la teoria de nombres: donat N, si existeix un enter x tal que $x^2 \equiv 1 \pmod{N}$ però $x \not\equiv \pm 1 \pmod{N}$, aleshores N és un enter compost i x s'anomena *arrel no trivial de* 1 mod N. El test per a trobar un testimoni és bàsicament una adaptació de l'algorisme indeterminista.

> Donat N, s'escriu $N-1$ com $2^k m$ amb m senar (la representació binària de $N-1$ és la de m seguida de k zeros). Escollim una *base* $a \in \{1, \ldots, N-1\}$ i considerem la seqüència $a^m, a^{2m}, \ldots, a^{2^k m} = a^{N-1}$.

Si per algun $2 \leq s \leq k$ obtenim $a^{2^s m} \equiv 1 \pmod{N}$ però $a^{2^{s-1} m} \not\equiv 1 \pmod{N}$, aleshores $a^{2^{s-1} m}$ és una arrel no trivial de 1 mod N i a és un testimoni que N és compost.

Si no trobem a la seqüència cap arrel no trivial de 1 mod N, aleshores comprovem si $a^{N-1} \not\equiv 1 \pmod{N}$. En cas afirmatiu, a és un testimoni que N és compost. En cas que no, a no és un testimoni i N podria ser primer (es poden trobar més detalls d'aquest algorisme a la secció 31.8 de [9], per exemple).

Quantes bases cal provar per estar segurs que N és compost? Miller [13] va donar un algorisme *determinista* per a comprovar la primalitat de N en temps polinòmic en $\log N$ si s'assumeix la veracitat de la Hipòtesi Estesa de Riemann (EHR en l'acrònim anglès). La Hipòtesi de Riemann estableix que els zeros positius de la funció de Riemann $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$, tenen part real $1/2$. La ERH és una extensió de conjectura de Riemann que tracta sobre la distribució de la part real dels zeros de la funció L de Dirichlet en el pla complex. Amb la hipòtesi EHR, Miller prova que , si N és compost, aleshores hi ha un testimoni amb $O(\log \log N)$ bits i, per tant, una cerca exhaustiva només ha de comprovar bases de llargada $\log \log N$, és a dir, un nombre polinòmic (en $\log N$) de bases. Així doncs, PRIMERS *pertany a la classe P sota la hipòtesis ERH.*

Rabin [17] (l'algorisme sense les demostracions va aparèixer primer a [16]) va proposar un mètode Monte-Carlo per decidir la primalitat d'un enter N. En lloc d'assumir l'EHR, l'algorisme selecciona aleatòriament amb una distribució uniforme una base $a \in \{1, \ldots, N-1\}$ i executa l'algorisme anterior per determinar si a és un testimoni per al caràcter compost de N. Si a no és un testimoni, aleshores o bé N pot ser primer o bé la tria aleatòria pot haver estat incorrecta. Rabin va demostrar que si $N > 4$ és compost, aleshores el nombre de testimonis de N és almenys $\frac{3(N-1)}{4}$ (la cota inferior es pot millorar a $\frac{N-1}{4}$ per a N senar i compost [5]). Repetint el procediment per a k testimonis (el que en termes algorísmics es coneix com *amplificant*), la probabilitat d'error és com a molt $\frac{1}{2^k}$, que es pot fer tan petita com es vulgui al preu d'augmentar el nombre k d'iteracions. Cal observar que l'algorisme de Rabin proporciona una comprovació en temps polinòmic (el que s'anomena *certificat curt*, o *short certificate*) del caràcter compost de N i no dóna, però, un certificat curt de primalitat en el cas que N és primer. Per tant, l'algorisme de Rabin posa el problema COMPOST a la classe RP (acrònim anglès de *Random polynomial time*) i, en conseqüència, PRIMERS pertany a la classe $co - RP$. La classe RP és la classe de problemes que poden ser resolts per un algorisme aleatori tipus Monte Carlo en temps polinòmic. Es coneix poc sobre aquesta classe; se sap que $RP \subseteq NP$, però, és un problema obert si $RP \subseteq NP \cap co - NP$. L'algorisme de Rabin és fàcil d'implementar i es fa servir de forma extensiva per generar enters grans que siguin *amb molt alta probabilitat* primers (vegeu per exemple [6]).

Paral·lelament a aquests desenvolupaments, Solovay i Strassen [18] van dissenyar un altre algorisme de temps polinòmic del tipus Monte-Carlo per com-

provar la primalitat d'un enter donat N. L'esquelet de l'algorisme és semblant a l'anterior: es tracta de seleccionar aleatòriament amb la distribució uniforme un element $a \in \{1,\ldots,N-1\}$ i comprovar si és un testimoni per al caràcter compost de N. En lloc de fer servir l'equació (1), però, el test que proposen està basat en el criteri d'Euler: donat un nombre primer N, un enter positiu $a < n$ amb $\text{mcd}(a,N) = 1$ és un residu quadràtic mòdul N (és a dir, $a = b^2 \pmod N$ per a algun b) si i només si $a^{\frac{N-1}{2}} \equiv 1 \pmod N$, on $a^{\frac{N-1}{2}} \pmod N$ s'anomena el *símbol de Legendre*, que es denota per $\left[\frac{a}{N}\right]$. El símbol de Legendre es pot calcular en temps polinòmic fent quadrats successius. El *símbol de Jacobi* n'és una generalització definida per

$$\left[\frac{a}{N}\right] = \prod_{k_i} \left[\frac{a}{p_i}\right]^{k_i},$$

on $N = p_1^{k_1} \cdots p_l^{k_l}$ és la factorització de N en nombres primers. El símbol de Jacobi es pot calcular de manera força eficient sense necessitat de conèixer els factors primers de N [7]. Solovay i Strassen van proposar el test següent: $a \in \{1,\ldots,N-1\}$ és un testimoni del caràcter compost de N si, o bé $\text{mcd}(a,N) > 1$, o bé el símbol de Jacobi $\left[\frac{a}{N}\right]$ i $a^{(N-1)}2 \pmod N$ no són congruents mod N. Aquest algorisme també proporciona un certificat curt del caràcter compost de N, demostrant que PRIMERS $\in co-RP$. A més, l'algorisme pot ser *desaleatoritzat* assumint l'ERH.

Durant la dècada dels vuitanta es va fer una feina considerable sobre el problema PRIMERS. Adleman, Pommerance i Rumely [3] van dissenyar un nou test de primalitat que corre en temps $O((\log N)^{c \log\log\log N})$, on c és una constant eficientment calculable. El seu mètode està inspirat en la idea dels residus quadràtics de Solovay i Strassen afegint-hi una versió més forta del Petit Teorema de Fermat que fa servir la teoria d'ideals. Cohen i Lenstra [7] van donar una versió modificada del test de primalitat d'Adleman, Pommerance i Rumely que és fàcilment programable per a aplicacions pràctiques. Un dels tests de primalitat més ràpids que existeix per a enters molt grans ve donat per la funció que es troba al sistema PARI i que fa servir el test de Cohen i Lenstra (`http://www.parigp-home.de`). Goldwasser i Kilian [11] van donar un algorisme aleatori polinòmic per PRIMERS sota la hipòtesi que, per a N gran, l'interval entre N i el següent nombre primer no és més llarg que $\log^2 N$ (en teoria de nombres aquesta hipòtesi es coneix com la conjectura de Cramér). A més, donat un primer N, el seu resultat proporciona un certificat curt de la primalitat de N. Així, suposant certa la conjectura de Cramér, aquest algorisme posa PRIMERS a la classe RP.

Finalment, Adleman i Huang [2] van modificar aquest resultat i van donar un algorisme tipus *Las Vegas* per PRIMERS sense cap hipòtesi addicional, de manera que PRIMERS $\in RP$, el que implica que PRIMERS $\in RP \cap co-RP = ZPP$ (ZPP és l'acrònim de *Zero probability error randomized polynomial time*, la classe de problemes reconeguts per algorismes aleatoris tipus Las Vegas, en

temps polinòmic). A l'excel·lent article panoràmic d'Adleman [1], hi ha una introducció clara als dos algorismes.

3 El present

La clau del nou algorisme AKS [4] està en la següent extensió del Petit Teorema de Fermat: donats dos enters a, n tals que $\mathrm{mcd}(a, N) = 1$, aleshores N és primer si i només si

$$(x - a)^N \equiv (x^N - a) \bmod N\,. \tag{2}$$

Per tant, donat N com a entrada, per executar un test de primalitat per a N escollim a amb $\mathrm{mcd}(a, N) = 1$ i verifiquem l'equació (2). El càlcul de $(x - a)^N$, però, és exponencial (en $\log N$) si hem de provar totes les a's possibles. Per superar aquest problema, els autors avalúen les dues bandes de l'equació (2) mòdul un polinomi $x^r - 1$ per a un valor adequat de r, és a dir, calculen

$$(x - a)^N \equiv (x^N - a) \bmod x^r - 1, N. \tag{3}$$

Aquí, adequat vol dir que r sigui un nombre primer i que $r - 1$ tingui un factor primer q gran que divideixi l'ordre de $N \pmod r$. L'existència d'un r amb aquesta propietat està garantida per resultats previs en teoria de nombres. A més, els autors proven que, per a N prou gran, un valor adequat de r es pot trobar a l'interval $[c_1(logN)^6, c_2(\log N)^6]$ per a certes constants positives c_1 i c_2, de manera que $r \sim \Theta((\log N)^6)$.

Així doncs, per executar un test de primalitat per a N donat, l'algorisme AKS consisteix en dues iteracions. En la primera, es selecciona un valor adequat de r. Començant en $r = 2$ i incrementant r en una unitat a cada pas, comproven que $\mathrm{mcd}(r, N) = 1$ (altrament N és compost), comproven que r és primer per cerca exhaustiva (que suposa comprovar $\Theta(logN)^6$ possibles factors), que el factor més gran q de $r - 1$ satisfà $q \geq 4(\log N)r^{1/2}$ i que $n^{(r-1)/q} \not\equiv 1 \pmod r$. Un cop seleccionat el valor de r, l'algorisme fa una segona iteració que verifica l'equació (3) per a tots els valors d'a entre 1 i $4(\log N)r^{1/2}$ (la tria de la mida en aquesta segona iteració es fa per raons tècniques en la demostració de la correcció de l'algorisme). Si en qualsevol iteració la equació (3) deixa de satisfer-se, aleshores N és compost. Altrament, N és primer.

La complexitat de l'algorisme AKS és $O((\log n)^{12} p(\log\log N))$, on p és un polinomi. La demostració que l'algorisme és correcte és relativament directa. Per més detalls, el lector pot mirar l'article d'AKS, que es força senzill d'entendre.

En el mateix article, els autors també proven que l'exponent 12 en l'expressió de complexitat de l'algorisme pot ser reduït a 6 si s'assumeix una conjectura sobre la densitat de la classe de nombres primers anomenats de Sophie Germain.

El que queda obert és l'elaboració d'enginyeria algorísmica per veure com es compara empíricament l'algorisme AKS amb les implementacions actuals

dels algorismes anteriors que es poden trobar com a funcions en sistemes com PARI, Mathematica o Maple.

Per acabar, voldria expressar el meu agraïment a O. Serra ja que sense ell aquest article mai s'hauria publicat i a A. Atserias, M. Agrawal i L. Fortnow pels seus valuosos comentaris a la versió apareguda al *Bulletin of the EATCS*.

Referències

[1] ADLEMAN, L. M. «Algorithmic number theory-the complexity contribution». *35th IEEE Symposium on Foundations of Computer Science*, (1994) 202–209.

[2] ADLEMAN, L. M.; HUANG, M.-D. «Recognizing primes in random polynomial time». *19th ACM Symposium on Theory of Computing*, (1987) 462–469.

[3] ADLEMAN, L. M.; POMERANCE, C.; RUMELY, R. S. «On distinguishing prime numbers from composite numbers». *Annals of Mathematics*, (1983) 117:173–206.

[4] AGRAWAL, M.; KAYAL, N.; SAXENA, N. «Primes in p». *43th IEEE Symposium on Foundations of Computer Science*, (2002) 302–309.

[5] BACH, E. *Analytic methods in the analysis and design of number-theoretic algorithms*. A.C.M. Distinguished Dissertations. The MIT Press, 1985.

[6] BEAUCHEMIN, P.; BRASSARD, G.; CREPEAU, C.; GOUTIER, C.; POMERANCE, C. «The generation of random numbers that are probably prime». *Journal of Criptology*, (1988) 1:53–64.

[7] COHEN, H.; LENSTRA, A. K. «Implementation of a new primality test». *Mathematics of computation*, (1984) 48:103–121.

[8] COOK, S. *The complexity of theorem-proving procedures*. 3rd. ACM Symposium on the Theory of Computing. ACM Press, 1971, 151–158.

[9] CORMEN, T. H.; LEISERSON, C.; RIVEST, R.; STEIN, C. *Introduction to Algorithms (Second Edition)*. Cambridge, Mass: MIT Press, 2001.

[10] GAREY, M. R.; JOHNSON, D. S. *Computers and tractability: A Guide to the Theory of NP-Completeness*. San Francisco: Freeman, 1979.

[11] GOLDWASSER, S.; KILIAN. J. «Almost all primes can be quickly certified». *18th ACM Symposium on Theory of Computing*, pages (1986) 316–329.

[12] KARP, R. M. «Reducibility among combinatorial problems». A: MILLER R. E., THATCHER J. W., ed. *Complexity of Computer Computations*, Plenum Press, NY, 1972, 85–104.

[13] MILLER, G. L. «Riemann's hypothesis and tests for primality». *Journal of Computer and System Sciences*, (1976) 13:300–317.

[14] PAPADIMITRIOU, C. *Computational Complexity*. Reading, Mass.: Addison-Wesley, 1994.

[15] PRATT, V. «Every prime has a succinct certificate». *SIAM Journal on Computing*, (1975) 4:214–220.

[16] Rabin, M. «Probabilistic algorithms». Traub, J. F. ed., *Algorithms and Complexity: New directions and recent results*. Academic Press, 1976, 21–39.

[17] Rabin, M. «Probabilistic algorithm for testing primality». *Journal of Number Theory*, (1980) 12:128–138.

[18] Solovay, R.; Strassen, V. «A fast Monte-Carlo test for primality». *SIAM Journal on Computing*, (1977) 6:84–86.

Departament de Llenguatges i Sistemes Informàtics
Universitat Politècnica de Catalunya
Campus Nord, Jordi Girona, 1–3, Edif. C5/6
E-08034, Barcelona
diaz@lsi.upc.es