INFORMATIQUE THÉORIQUE ET APPLICATIONS

JACQUES ARSAC

Algorithmes pour vérifier la conjecture de Syracuse

Informatique théorique et applications, tome 21, nº 1 (1987), p. 3-9.
<http://www.numdam.org/item?id=ITA_1987__21_1_3_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 21, n° 1, 1987, p. 3 à 9)

ALGORITHMES POUR VÉRIFIER LA CONJECTURE DE SYRACUSE (*)

par Jacques ARSAC ([1])

Communiqué par J. BERSTEL

Résumé. – *La suite de Syracuse de l'entier naturel n est définie par*

$$U(n, 0) = n.$$

$$U(n, i+1) = SI \text{ pair } U(n, i) \text{ } ALORS \text{ } U(n, i)/2 \text{ } SINON \text{ } (3\,U(n, i)+1)/2.$$

La conjecture de Syracuse dit que pour tout $n \in \mathbb{N}$, $n > 1$, il existe k fini tel que $U(n, k) = 1$. Ceci n'a pu être démontré. On a vérifié que la conjecture est vraie jusqu'à 2^{40} par calculs sur ordinateur. On examine ici la complexité de ces calculs, et on montre qu'elle peut être réduite grâce aux propriétés des suites de Syracuse.

Abstract. – *The Syracuse Sequence of natural integer n, n is defined as*

$$U(n, 0) = n$$

$$U(n, i+1) = IF \text{ even } U(n, i) \text{ } THEN \text{ } U(n, i)/2 \text{ } ELSE \text{ } (3\,U(n, i)+1)/2$$

It has been conjectured that for every $n \in \mathbb{N}$, $n > 1$ there exists a finite integer k such that $U(n, k) = 1$. This fact has not yet been proved. The conjecture has been verified on a computer up to 2^{40}. We consider here the complexity of direct algorithms to check the conjecture, and the way in which the complexity may be reduced through convenient properties of sequences of Syracuse.

1. INTRODUCTION

Dans ce qui suit, tous les nombres sont des entiers naturels. La suite de Syracuse de $n > 1$ est définie par

$$U(n, 0) = n$$

$$U(n, i+1) = \text{Si pair } U(n, i) \text{ ALORS } U(n, i)/2 \text{ SINON } (3\,U(n, i)+1)/2$$

(*) Reçu en décembre 1984, révisé en juin 1986.
([1]) Université Pierre-et-Marie-Curie, 4, place Jussieu, 75005 Paris.

La conjecture de Syracuse dit que pour tout $n>1$ il existe $k(n)$ fini tel que $U(n, k(n))=1$. Ceci n'a pas été encore démontré, et Lagarias [1] dit ce problème non traitable, bien que l'on possède quelques théorèmes. Il demande s'il existe de bons programmes pour vérifier la conjecture sur ordinateur jusqu'à un entier N aussi grand que possible. Dans une lettre à J. C. Lagarias, Shiro Ando [2] annonce qu'il l'a fait jusqu'à 2 ^ 40. Le programme utilisé n'ayant pas été publié, on peut penser qu'il était simple. Nous nous proposons de montrer ici comment les algorithmes simples peuvent être améliorés.

2. ALGORITHMES SIMPLES

Pour vérifier la conjecture, il suffit de calculer les suites de Syracuse des entiers n entre 2 et N jusqu'au plus petit i pour lequel $U(n, i)=1$ (appelé « temps d'arrêt total de n » [1]).

```
POUR n:=2 JUSQUA N FAIRE
  r:=n
  TANT QUE r>1 FAIRE
    SI pair r ALORS r:=r/2 SINON r:=(3r+1)/2 IS
  BOUCLER
BOUCLER
```

On conjecture que le temps d'arrêt total de n est borné par $C \log(n)$. Le nombre de pas exécutés par ce programme serait alors borné par une quantité en $N \log(N)$. L'expérience faite sur l'intervalle

$$1\,024<N<8\,192$$

donne un nombre de pas croissant très légèrement plus vite que $N \log(N)$.

Mais on fait beaucoup trop de calculs. Par exemple $U(9,2)=7$. A partir de là, la suite de Syracuse de 9 coïncide avec celle de 7, dont on sait qu'elle arrive à 1 parce que 7 a déjà été testé dans le programme. Il suffit donc d'arrêter la suite de n au plus petit i tel que $U(n, i)<n$, appelé « temps d'arrêt simple de n » dans [1]. On change le test $r>1$ en $r \geqq n$ dans le programme. L'expérience faite dans les mêmes conditions donne un nombre de pas très proche de 3.5 N, ce qui est une amélioration considérable.

3. NOMBRES DONT LE TEMPS D'ARRÊT EST CONNU

Si un nombre a un temps d'arrêt simple connu *a priori*, il est inutile de le tester. Ainsi

$U(2n', 1)=n'<2n'$. Tout nombre pair a 1 pour temps d'arrêt.

$U(4n'+1, 2)=3n'+1<4n'+1$: temps d'arrêt 2.

On peut ne pas tester les nombres pairs, ce qui supprime $N/2$ nombres mais seulement $N/2$ pas, ni les nombres congrus à 1 modulo 4, ce qui enlève $N/4$ nombres, et 2 pas pour chacun, soit encore $N/2$ pas. Mais la complexité reste $O(N)$.

Appelons «montée» un pas qui fait passer de r à $(3r+1)/2$. Nous dirons que n donne r en m pas dont a montées, et nous noterons

$$n(m, a)r$$

le fait que $U(n, m)=r$ avec a montées dans le calcul de $U(n, m)$.

THÉORÈME 1: *Soit n $(m, a)r$. Pour tout $k \in \mathbb{N}$, $n+k.2\wedge m$ $(m, a)r+k.3\wedge a$ par la même suite de m pas. En outre*

$$(r/n)>(r+k.3\wedge a)/(n+k.2\wedge m)>3\wedge a/2\wedge m$$

La première partie se démontre par récurrence sur m. Pour la deuxième partie, on utilise le fait que

$$r=(3\wedge a n+x(n))/2\wedge m \text{ pour un certain } x(n)>0.$$

Comme corollaire, on retrouve le fait que $x(n+2\wedge m)=x(n)$: x est périodique (théorème dû à Terras [3]).

Si n a un temps d'arrêt simple m, $r<n$ et donc $3\wedge a<2\wedge m$. On a cherché une réciproque à ce résultat, mais il y a des contre-exemples

$$3\wedge 5=243<2\wedge 8=256 \quad \text{or } 7\ (8,5)\ 8>7.$$

Si n a un temps d'arrêt simple m, avec a montées, alors pour tout entier naturel k

$$n+k.2\wedge m\ (m, a)\ r+k.3\wedge a<n+k.2\wedge m.$$

Ainsi, tous les entiers congrus à n modulo $2\wedge m$ ont le même temps d'arrêt et n'ont pas à être testés.

THÉORÈME 2: *Pour toute suite s de pas de longueur m commençant par une montée, il existe n unique inférieur à $2\wedge m$ ayant s comme m premiers pas de sa suite de Syracuse.*

On le démontre en donnant l'algorithme construisant n. Supposons que l'on ait obtenu $n(i)$ dont la suite de Syracuse commence par les i premiers pas de s comportant $a(i)$ montées:

$$n(i)\ (i, a(i))\ r(i) \qquad \text{avec} \quad n(i) < 2 \mathbin{\char`\^} i$$

4 cas sont possibles:

1. $r(i)$ est pair, le pas suivant de s est descendant:

$$n(i+1) = n(i)\ n(i)\ (i+1, a(i))\, r(i)/2 \qquad \text{et} \qquad n(i) < 2 \mathbin{\char`\^} i$$

2. $r(i)$ est impair et le pas suivant ascendant. On a encore $n(i+1) = n(i)$.

Dans les deux autres cas, $r(i)$ n'a pas la bonne parité. Mais

$$n(i) + 2 \mathbin{\char`\^} i\, (i, a(i))\, r(i) + 3 \mathbin{\char`\^} a(i)$$

donne la parité opposée, et on prend

$$n(i+1) = n(i) + 2 \mathbin{\char`\^} i < 2 \mathbin{\char`\^} (i+1)$$

On peut démarrer avec $n(1) = 1\ r(1) = 2$.

Comme corollaire, les m premiers pas de la suite de n sont donnés par les m chiffres de droite dans l'écriture de n en binaire. Ce résultat était connu [1], mais apparemment pas l'algorithme de calcul de n.

4. UN CRIBLE

Si, dans la suite des nombres à tester, on trouve n tel que

$$n\,(m, a)\ r < n$$

alors on peut rayer tous les nombres à partir de n au pas $2 \mathbin{\char`\^} m$. Dans le cas contraire, supposons que l'on ait en outre $3 \mathbin{\char`\^} a > 2 \mathbin{\char`\^} m$ (qui implique $r > n$). Alors par le théorème 1, pour tout k

$$n + k \,.\, 2 \mathbin{\char`\^} m\ (m, a)\ r + k \,.\, 3 \mathbin{\char`\^} a > n + k \,.\, 2 \mathbin{\char`\^} m$$

Ainsi, aucun nombre congru à r modulo $3 \mathbin{\char`\^} a$ n'a à être testé, car il est dans la suite de Syracuse d'un nombre plus petit, déjà testé, et dont on sait que son temps d'arrêt simple est fini. Par exemple

$2k+1\ (1, 1)\ 3k+2$: aucun nombre égal à 2 modulo 3 n'a à être testé.

$3\,(3, 2)\ 4$: aucun nombre égal à 4 modulo 9 n'a à être testé.

Mais ce crible est fort limité. Aucun terme dans la suite de Syracuse d'un nombre impair ne peut être congru à 0 modulo 3. On pourra rayer au plus

2/3 des nombres. En outre, des nombres différents peuvent donner, le même résultat. Par exemple

7 (4, 3) 13 mais les nombres égaux à 13 modulo 27 sont égaux à 4 modulo 9. Nous avons construit successivement la table des nombres rayables modulo 2 187, puis 6 561 : ce sont les mêmes. Nous nous en sommes tenus à 2 187, qui amène à rayer 46 % des nombres.

On ne peut réaliser le crible sur odinateur en tenant à jour la liste des nombres à tester et en rayant dedans, parce que cela demande trop de place en mémoire (même si chaque nombre est représenté par un seul bit disant s'il est rayé ou non).

Soit $F(n, m, a, r, h)$ la procédure qui teste la conjecture de Syracuse pour les entiers $n+k.2\hat{}\ m$ jusqu'à $2\hat{}\ h$, $k \geqq 0$. La conjecture sera vérifiée pour tous les entiers jusqu'à $2\hat{}\ h$ par $F(3, 2, 2, 8, h)$. Si $r<n$, par le théorème 1, la conjecture est vraie sur l'ensemble considéré et rien n'est à faire. Si $m=h$, seul n est dans l'intervalle étudié. S'il ne peut être éliminé à cause de son reste dans la division par 2 187, il faut vérifier son temps d'arrêt simple. Si enfin $m<h$, on décompose le problème en deux :

r pair :

$$n(m+1, a)\ r/2, \qquad n+2\hat{}\ m\ (m+1,\ a+1)\ (3(r+3\hat{}\ a)+1)/2$$

r impair :

$$n(m+1, a+1)\ (3r+1)/2, \qquad n+2\hat{}\ m\ (m+1, a)\ (r+3\hat{}\ a)/2.$$

Ceci donne une procédure F récursive dyadique. Nous n'avons aucune façon d'en déterminer la complexité par la théorie. L'expérience faite en faisant varier h de 8 à 19 a donné les résultats suivants :

nombre d'entiers dont l'arrêt doit être vérifié $=(N\hat{}\ 0{,}87)/16$;

nombre de pas pour ces vérifications $=0{,}88\ (N\hat{}\ 0{,}87)$.

5. NOMBRES PRIMITIFS

L'expérience montre que plusieurs nombres peuvent donner à partir d'un certain rang la même suite de Syracuse, provoquant des calculs inutiles :

$$U(31{,}8)=91, \qquad U(63{,}9)=91$$

bien que 63 (multiple de 3) ne soit pas dans la suite de 31. On appelle « primitif d'ordre p » tout nombre dont la suite de Syracuse commence par

$p>0$ montées suivieis d'une seule descente. Si les p premières montées sont suivies d'au moins 2 descentes, n est dit « non primitif d'ordre p ».

THÉORÈME 3: *Si n est primitif d'ordre p, alors $n+k.2\wedge(p+2)$ est aussi primitif d'ordre p, $2n+1+k.2\wedge(p+3)$ est non primatif d'ordre $p+1$ pour tout $k \geqq 0$. En outre $U(2n+1, p+3)=U(n, p+2)$.*

Quel que soit n impair, $4n+1$ est non primatif, et $U(4n+1,3)=U(n, 1)$.

Le fait que la primitivité ou non-primitivité soit conservée modulo une puissance de 2 résulte de la définition. On montre aisément que tout nombre primitif d'ordre p est de la forme $t.2\wedge p-1$ (t strictement positif) de sorte que $n=t.2\wedge p-1$ donne en p montées suivies d'une descente $r=t.3\wedge p-1$. $2n+1=t.2\wedge(p+1)-1$ donne en $p+1$ montées suivies d'une descente $t.3\wedge(p+1)-1=(3r+1)$ pair. $2n+1$ est donc non primitif. Le pas suivant donne $(3r+1)/2$, terme suivant de la suite de r, donc de n.

Par ce théorème, aucun nombre non primitif n'a à être testé, car son arrêt se déduit de $(n-1)/2$ ou $(n-1)/4$, plus petit. Quand l'évaluation de la suite de Syracuse de n donne au début p montées suivies de 2 descentes, alors n est non primitif d'ordre p, on peut arrêter l'évaluation de la suite et rayer les nombres à partir de n au pas $p+2$.

Soit maintenant $n(m, a)r$ non primitif d'ordre p. Son arrêt se déduit de celui de $(r-1)/2$ si $p>1$, $(r-1)/4$ si $p=1$. Si ces valeurs sont inférieures à n, on peut arrêter la vérification de l'arrêt de n. Par le théorème 1, ceci entraîne $3\wedge a<2\wedge(m+1)$ pour $p>1$ et $3\wedge a<2\wedge(m+2)$ si $p=1$. Supposons au contraire que

$$n(m, a)\ r \text{ non primitif d'ordre } p>1$$

et $3\wedge a>2\wedge(m+1)$ entrainant $(r-1)/2>n$. Soit s le plus petit entier tel que $2s+1=r$ modulo $3\wedge a$. Aucun entier congru à s modulo $3\wedge a$ n'a à être essayé. Soit en effet $q=s+k.3\wedge a$ un tel entier. Ou bien il est non primitif, et n'a pas à être essayé. S'il n'en est pas ainsi, il est primitif, mais alors

$2q+1=2s+1+2k.3\wedge a=r+k'.3\wedge a$ est non primitif et a même arrêt. Or ce nombre est dans la suite de Syracuse de

$$n+k'.2\wedge \mathrm{m}\ (m,\ a)\ r+k'.3\wedge a$$

et $(r+k'.3\wedge a-1)/2=q>n+k'.2\wedge m$. Ainsi l'arrêt de q se déduit de celui d'un nombre inférieur à q. On établit un résultat analogue pour les non primitifs d'ordre 1.

On complète ainsi le crible présenté précédemment. On accroît le nombre d'entiers que l'on peut rayer modulo une puissance de 3 (on a conservé

$3\char`^ 7=2\,187$). Le crible sur les puissances de 2 est lui aussi étendu par la détection des nombres non primitifs. Ceci réduit le temps de calcul, mais ne change pas fondamentalement la complexité

nombre d'entiers dont l'arrêt doit être vérifié: $(N\char`^ 0{,}866)/19{,}4$;

nombre de pas pour ces vérifications: $0{,}73\ (N\char`^ 0{,}866)$.

Il y a un très léger gain sur la puissance de N, et un gain sur le coefficient. Notons que l'on peut construire un algorithme incrémental qui vérifie la conjecture jusqu'à $2\char`^ h$ en la supposant vraie jusqu'à $2\char`^ (h-1)$ (ceci permet de passer d'abord à $2\char`^ 41$, puis $2\char`^ 42\ldots$). C'est une modification facile des programmes précédents, réduisant encore le coût des calculs. Mais il faut aussi voir que r croît plus vite que n, ce qui oblige à conduire les calculs en multiple précision. On peut toutefois mentionner le théorème suivant:

THÉORÈME 4: *Soit* $n(m, a)\ r$. *Si* $n<2\char`^ m$, *alors* $r<3\char`^ a$.

On le démontre par récurrence sur m, en vérifiant que le théorème est vrai, par exemple pour $m=3$. Ce résultat est un peu meilleur que ce que l'on peut déduire des résultats antérieurs [1], par exemple de

$$r=(3\char`^ a\ n+x)/2\char`^ m, \qquad x<2\char`^ (m-a)\ (3\char`^ a-2\char`^ a).$$

Ceci vient de ce qu'il y a une relation entre n et la suite qu'il engendre: on ne peut se donner séparément la suite m, a et n. En particulier, si $3\char`^ a<2\char`^ m$, et s'il existe n: $3\char`^ a<n<2\char`^ m$ donnant r en une suite m, a

$$3\char`^ a<n<2\char`^ m\ (m,\ a)\ r<3\char`^ a<n$$

C'est un cas où la réciproque mentionnée pour le théorème 1 est vraie.

BIBLIOGRAPHIE

1. J. C. LAGARIAS, *The* $3x+1$ *Problem and its Generalizations*, American Monthly, janvier 1985, p. 3-21.
2. S. ANDO, Letter to J. C. Lagarias.
3. R. TERRAS *A Stopping Time Problem on the Positive Integers*, Acta arithmetica, vol. 30, 1976, p. 241-252.