Algorithmique Probabiliste

Philippe Duchon

LaBRI - ENSEIRB-Matméca - Université de Bordeaux

2014-15

Algorithmes probabilistes en arithmétique

- **Test de Miller-Rabin :** pour savoir si un entier donné est premier ou non

Algorithmes probabilistes en arithmétique

- **Test de Miller-Rabin :** pour savoir si un entier donné est premier ou non
- **Application :** générer aléatoirement de grands nombres premiers

Algorithmes probabilistes en arithmétique

- **Test de Miller-Rabin :** pour savoir si un entier donné est premier ou non
- **Application :** générer aléatoirement de grands nombres premiers
- (**Racines carrées modulaires :** étant donnés deux entiers x et n, trouver, s'il en existe, une "racine carrée de x modulo n" (y tel que $y^2 = x \mod n$, cas où n est premier))

Algorithmes probabilistes en arithmétique

- **Test de Miller-Rabin :** pour savoir si un entier donné est premier ou non
- **Application :** générer aléatoirement de grands nombres premiers
- (**Racines carrées modulaires :** étant donnés deux entiers x et n, trouver, s'il en existe, une "racine carrée de x modulo n" (y tel que $y^2 = x \mod n$, cas où n est premier))
- Tout cela a des applications en cryptologie : énormément de cryptosystèmes sont basés sur de la théorie des nombres

Quelques faits sur les nombres premiers

- **nombre premier :** entier > 1 dont les seuls diviseurs entiers sont 1 et lui-même

Quelques faits sur les nombres premiers

- **nombre premier :** entier > 1 dont les seuls diviseurs entiers sont 1 et lui-même
- l'ensemble $\mathcal{P}$ des nombres premiers est *infini* (c'est connu depuis l'Antiquité)

Quelques faits sur les nombres premiers

- **nombre premier :** entier > 1 dont les seuls diviseurs entiers sont 1 et lui-même
- l'ensemble $\mathcal{P}$ des nombres premiers est *infini* (c'est connu depuis l'Antiquité)
- les grands nombres premiers sont “rares”, mais “pas tant que ça” : si $\pi(n)$ désigne le nombre de nombres premiers inférieurs à n, on a (**Théorème des nombres premiers**)

$$\pi(n) \simeq \frac{n}{\ln n}$$

  donc en particulier le nombre d'entiers premiers entre 2^k et 2^{k+1} est d'ordre $2^{k+1}\ln(2)/(k+1) - 2^k \ln(2)/k \simeq 2^k \ln(2)/k$ (une proportion de l'ordre de $0.7/k$).

Quelques faits sur les nombres premiers

- **nombre premier :** entier > 1 dont les seuls diviseurs entiers sont 1 et lui-même
- l'ensemble $\mathcal{P}$ des nombres premiers est *infini* (c'est connu depuis l'Antiquité)
- les grands nombres premiers sont "rares", mais "pas tant que ça" : si $\pi(n)$ désigne le nombre de nombres premiers inférieurs à n, on a (**Théorème des nombres premiers**)

$$\pi(n) \simeq \frac{n}{\ln n}$$

  donc en particulier le nombre d'entiers premiers entre 2^k et 2^{k+1} est d'ordre $2^{k+1} \ln(2)/(k+1) - 2^k \ln(2)/k \simeq 2^k \ln(2)/k$ (une proportion de l'ordre de $0.7/k$).
- **(petit) théorème de Fermat :** si n est un nombre premier, alors pour tout entier $a \in [[1, n-1]]$, $a^{n-1} = 1 \mod n$

Un peu d'arithmétique

- Si n est premier, alors $\mathbb{Z}_n$ (entiers modulo n, avec addition et multiplication) est un corps (tout entier autre que 0 est inversible) **et réciproquement**

Un peu d'arithmétique

- Si n est premier, alors $\mathbb{Z}_n$ (entiers modulo n, avec addition et multiplication) est un corps (tout entier autre que 0 est inversible) **et réciproquement**
- Connaissant n (même sans savoir si n est premier, même sans connaître la factorisation de n) les calculs dans $\mathbb{Z}_n$ se font bien : addition, multiplication, *exponentiation modulaire* ($a^b \bmod n$), *calcul d'inverse* (par PGCD)

Un peu d'arithmétique

- Si n est premier, alors $\mathbb{Z}_n$ (entiers modulo n, avec addition et multiplication) est un corps (tout entier autre que 0 est inversible) **et réciproquement**
- Connaissant n (même sans savoir si n est premier, même sans connaître la factorisation de n) les calculs dans $\mathbb{Z}_n$ se font bien : addition, multiplication, *exponentiation modulaire* ($a^b \bmod n$), *calcul d'inverse* (par PGCD)
- **En revanche**, on ne connaît aucun algorithme efficace pour factoriser un entier en produit de nombres premiers

Le “test de Fermat”

- Conséquence du petit théorème de Fermat : si on prend un entier a, qu'on calcule a^{n-1} modulo n, et que ça ne fait pas 1, alors on **sait** que n n'est pas premier : a est un “témoin de non primalité” pour n

Le “test de Fermat”

- Conséquence du petit théorème de Fermat : si on prend un entier a, qu'on calcule a^{n-1} modulo n, et que ça ne fait pas 1, alors on **sait** que n n'est pas premier : a est un “témoin de non primalité” pour n
- Algorithme possible : tirer un a aléatoire entre 1 et $n-1$, tester s'il s'agit d'un témoin de non primalité ; si c'est le cas, déclarer n non premier, sinon, déclarer n premier

Le “test de Fermat”

- Conséquence du petit théorème de Fermat : si on prend un entier a, qu'on calcule a^{n-1} modulo n, et que ça ne fait pas 1, alors on **sait** que n n'est pas premier : a est un “témoin de non primalité” pour n
- Algorithme possible : tirer un a aléatoire entre 1 et $n-1$, tester s'il s'agit d'un témoin de non primalité ; si c'est le cas, déclarer n non premier, sinon, déclarer n premier
- C'est un algorithme à erreur unilatérale : on ne peut se tromper que dans un seul sens, en déclarant n premier alors qu'il ne l'est pas (algorithme 1MC pour le problème “n est-il non premier ?”)

Le “test de Fermat”

- Conséquence du petit théorème de Fermat : si on prend un entier a, qu'on calcule a^{n-1} modulo n, et que ça ne fait pas 1, alors on **sait** que n n'est pas premier : a est un “témoin de non primalité” pour n
- Algorithme possible : tirer un a aléatoire entre 1 et $n-1$, tester s'il s'agit d'un témoin de non primalité ; si c'est le cas, déclarer n non premier, sinon, déclarer n premier
- C'est un algorithme à erreur unilatérale : on ne peut se tromper que dans un seul sens, en déclarant n premier alors qu'il ne l'est pas (algorithme 1MC pour le problème “n est-il non premier ?”)
- La question critique est celle de la probabilité d'erreur

Témoins pour le test de Fermat

- Si n est premier (n'est pas "non premier"), il n'a **aucun** témoin de non primalité (petit théorème de Fermat), la probabilité d'erreur est bien 0

Témoins pour le test de Fermat

- Si n est premier (n'est pas "non premier"), il n'a **aucun** témoin de non primalité (petit théorème de Fermat), la probabilité d'erreur est bien 0
- Si n n'est pas premier, il a au moins un diviseur premier $p < n$

Témoins pour le test de Fermat

- Si n est premier (n'est pas “non premier”), il n'a **aucun** témoin de non primalité (petit théorème de Fermat), la probabilité d'erreur est bien 0
- Si n n'est pas premier, il a au moins un diviseur premier $p < n$
- Dans ce cas, tous les multiples de p sont non premiers avec n, et donc non inversibles modulo n : leurs puissances ne peuvent être congrues à 1 modulo n - **ce sont des témoins de non primalité**

Témoins pour le test de Fermat

- Si n est premier (n'est pas "non premier"), il n'a **aucun** témoin de non primalité (petit théorème de Fermat), la probabilité d'erreur est bien 0
- Si n n'est pas premier, il a au moins un diviseur premier $p < n$
- Dans ce cas, tous les multiples de p sont non premiers avec n, et donc non inversibles modulo n : leurs puissances ne peuvent être congrues à 1 modulo n - **ce sont des témoins de non primalité**
- **Malheureusement,** il existe une infinité de nombres ("de Carmichael") qui sont tous *non premiers* et pour lesquels **les seuls témoins de non primalité sont les entiers non premiers avec eux**

Témoins pour le test de Fermat

- Si n est premier (n'est pas "non premier"), il n'a **aucun** témoin de non primalité (petit théorème de Fermat), la probabilité d'erreur est bien 0
- Si n n'est pas premier, il a au moins un diviseur premier $p < n$
- Dans ce cas, tous les multiples de p sont non premiers avec n, et donc non inversibles modulo n : leurs puissances ne peuvent être congrues à 1 modulo n - **ce sont des témoins de non primalité**
- **Malheureusement,** il existe une infinité de nombres ("de Carmichael") qui sont tous *non premiers* et pour lesquels **les seuls témoins de non primalité sont les entiers non premiers avec eux**
- 561, 1105, 1729, 2465...

Témoins pour le test de Fermat

- Si n est premier (n'est pas "non premier"), il n'a **aucun** témoin de non primalité (petit théorème de Fermat), la probabilité d'erreur est bien 0
- Si n n'est pas premier, il a au moins un diviseur premier $p < n$
- Dans ce cas, tous les multiples de p sont non premiers avec n, et donc non inversibles modulo n : leurs puissances ne peuvent être congrues à 1 modulo n - **ce sont des témoins de non primalité**
- **Malheureusement,** il existe une infinité de nombres ("de Carmichael") qui sont tous *non premiers* et pour lesquels **les seuls témoins de non primalité sont les entiers non premiers avec eux**
- 561, 1105, 1729, 2465...
- Si n est un nombre de Carmichael, la probabilité d'erreur du test de Fermat peut être *très proche de* 1 : si $n = p_1 \times \cdots \times p_k$, la proportion de témoins est de l'ordre de
$$\frac{1}{p_1} + \cdots + \frac{1}{p_k}$$

Le test de Miller-Rabin

- Le nombre de répétitions qu'il faudrait faire du test de Fermat pour abaisser la probabilité d'erreur à $1/2$, grandit presque comme n : c'est prohibitif (exponentiel en $\log(n)$).

Le test de Miller-Rabin

- Le nombre de répétitions qu'il faudrait faire du test de Fermat pour abaisser la probabilité d'erreur à $1/2$, grandit presque comme n : c'est prohibitif (exponentiel en $\log(n)$).
- Le **test de Miller-Rabin** est une forme de généralisation du test de Fermat : on augmente le nombre de témoins

Le test de Miller-Rabin

- Le nombre de répétitions qu'il faudrait faire du test de Fermat pour abaisser la probabilité d'erreur à $1/2$, grandit presque comme n : c'est prohibitif (exponentiel en $\log(n)$).
- Le **test de Miller-Rabin** est une forme de généralisation du test de Fermat : on augmente le nombre de témoins
- **Proposition :** si n est premier, alors 1 n'a que deux racines carrées modulo n (ce sont 1 et $n-1$) ; si n n'est pas premier, il en a plus. (Dans n'importe quel corps, une équation polynomiale comme $X^2 - 1 = 0$ a au plus un nombre de solutions égal à son degré ; inversement, le théoreme des restes chinois nous garantit que, vu que l'équation $X^2 - 1 = 0$ a deux solutions modulo p et deux solutions modulo q, elle en a quatre modulo $n = pq$)

Le test de Miller-Rabin

- Le nombre de répétitions qu'il faudrait faire du test de Fermat pour abaisser la probabilité d'erreur à $1/2$, grandit presque comme n : c'est prohibitif (exponentiel en $\log(n)$).
- Le **test de Miller-Rabin** est une forme de généralisation du test de Fermat : on augmente le nombre de témoins
- **Proposition :** si n est premier, alors 1 n'a que deux racines carrées modulo n (ce sont 1 et $n-1$) ; si n n'est pas premier, il en a plus. (Dans n'importe quel corps, une équation polynomiale comme $X^2 - 1 = 0$ a au plus un nombre de solutions égal à son degré ; inversement, le théoreme des restes chinois nous garantit que, vu que l'équation $X^2 - 1 = 0$ a deux solutions modulo p et deux solutions modulo q, elle en a quatre modulo $n = pq$)
- Idée du test de Miller-Rabin : un nombre a est un **témoin (fort) de non primalité** pour n si, soit c'est un témoin au sens de Fermat ($a^{n-1} \neq 1 \mod n$), soit il permet de trouver une racine carrée de 1 qui ne soit ni 1 ni $n-1$

Témoins forts de non primalité

- $n - 1$ est pair : on factorise $n - 1 = 2^s.r$, où r est impair

Témoins forts de non primalité

- $n - 1$ est pair : on factorise $n - 1 = 2^s.r$, où r est impair
- On calcule $a^r \mod n$, puis par élévations successives au carré, $a^{2r}, a^{4r}, \ldots, a^{2^s r}$

Témoins forts de non primalité

- $n-1$ est pair : on factorise $n-1=2^s.r$, où r est impair
- On calcule $a^r \mod n$, puis par élévations successives au carré, $a^{2r}, a^{4r}, \dots, a^{2^s r}$
- Si cette suite ne se termine pas par 1, a est un témoin (de Fermat) de non primalité de n

Témoins forts de non primalité

- $n - 1$ est pair : on factorise $n - 1 = 2^s.r$, où r est impair
- On calcule $a^r \mod n$, puis par élévations successives au carré, $a^{2r}, a^{4r}, \ldots, a^{2^s r}$
- Si cette suite ne se termine pas par 1, a est un témoin (de Fermat) de non primalité de n
- Si la suite se termine par 1, mais que le dernier terme qui précède le premier 1 est autre chose que $n - 1$, on a une racine carrée non triviale de 1 et donc n ne peut pas être premier (a est un "témoin fort de non primalité" pour n)

Témoins forts de non primalité

- $n - 1$ est pair : on factorise $n - 1 = 2^s.r$, où r est impair
- On calcule $a^r \mod n$, puis par élévations successives au carré, $a^{2r}, a^{4r}, \ldots, a^{2^s r}$
- Si cette suite ne se termine pas par 1, a est un témoin (de Fermat) de non primalité de n
- Si la suite se termine par 1, mais que le dernier terme qui précède le premier 1 est autre chose que $n - 1$, on a une racine carrée non triviale de 1 et donc n ne peut pas être premier (a est un "témoin fort de non primalité" pour n)
- Si la suite commence par 1, ou contient un $n - 1$ avant le premier 1, alors a n'est pas un témoin fort de non primalité.

Témoins forts de non primalité

- $n - 1$ est pair : on factorise $n - 1 = 2^s.r$, où r est impair
- On calcule $a^r \mod n$, puis par élévations successives au carré, $a^{2r}, a^{4r}, \ldots, a^{2^s r}$
- Si cette suite ne se termine pas par 1, a est un témoin (de Fermat) de non primalité de n
- Si la suite se termine par 1, mais que le dernier terme qui précède le premier 1 est autre chose que $n - 1$, on a une racine carrée non triviale de 1 et donc n ne peut pas être premier (a est un "témoin fort de non primalité" pour n)
- Si la suite commence par 1, ou contient un $n - 1$ avant le premier 1, alors a n'est pas un témoin fort de non primalité.
- C'est le test de Miller-Rabin !

```
IsWitness(a,n):
    a=a%n
    if a==0 or a==1:
        return False
    r=n-1
    s=0
    while r%2==0:
        r = r/2
        s = s+1
    x = PowMod(a,r,n)
    k = 0
    y=n-1
    while (k<s) and (x!=1):
        y = x
        x = (x*x) % n
        k = k+1
    if y!=n-1 or x!=1:
        return True
    return False
```

Proba d'erreur : nombre de témoins

- Le test de Miller-Rabin admet *plus* de témoins que celui de Fermat : sa probabilité d'erreur ne peut pas être pire

Proba d'erreur : nombre de témoins

- Le test de Miller-Rabin admet *plus* de témoins que celui de Fermat : sa probabilité d'erreur ne peut pas être pire
- Il faut tout de même une garantie : *minorer le nombre de témoins* pour un nombre non premier

Proba d'erreur : nombre de témoins

- Le test de Miller-Rabin admet *plus* de témoins que celui de Fermat : sa probabilité d'erreur ne peut pas être pire
- Il faut tout de même une garantie : *minorer le nombre de témoins* pour un nombre non premier
- Il se trouve que **le nombre de témoins (forts) de non primalité** pour un entier n non premier, est toujours **au moins** $3(n-1)/4$

Proba d'erreur : nombre de témoins

- Le test de Miller-Rabin admet *plus* de témoins que celui de Fermat : sa probabilité d'erreur ne peut pas être pire
- Il faut tout de même une garantie : *minorer le nombre de témoins* pour un nombre non premier
- Il se trouve que **le nombre de témoins (forts) de non primalité** pour un entier n non premier, est toujours **au moins** $3(n-1)/4$
- **Donc** la probabilité d'erreur de l'algorithme de Miller-Rabin, est toujours bornée par $1/4$

Proba d'erreur : nombre de témoins

- Le test de Miller-Rabin admet *plus* de témoins que celui de Fermat : sa probabilité d'erreur ne peut pas être pire
- Il faut tout de même une garantie : *minorer le nombre de témoins* pour un nombre non premier
- Il se trouve que **le nombre de témoins (forts) de non primalité** pour un entier n non premier, est toujours **au moins** $3(n-1)/4$
- **Donc** la probabilité d'erreur de l'algorithme de Miller-Rabin, est toujours bornée par $1/4$
- Comme pour tout algorithme à erreur unilatérale, en répétant k fois l'algorithme on abaisse cette probabilité d'erreur à $1/4^k$.

Proba d'erreur : nombre de témoins

- Le test de Miller-Rabin admet *plus* de témoins que celui de Fermat : sa probabilité d'erreur ne peut pas être pire
- Il faut tout de même une garantie : *minorer le nombre de témoins* pour un nombre non premier
- Il se trouve que **le nombre de témoins (forts) de non primalité** pour un entier n non premier, est toujours **au moins** $3(n-1)/4$
- **Donc** la probabilité d'erreur de l'algorithme de Miller-Rabin, est toujours bornée par $1/4$
- Comme pour tout algorithme à erreur unilatérale, en répétant k fois l'algorithme on abaisse cette probabilité d'erreur à $1/4^k$.
- (dans la pratique, la grande majorité des nombres non premiers ont beaucoup plus de témoins que ça ; c'est juste pour être sûr de majorer)

Tirage d'entiers (presque) premiers aléatoires

- **Le problème :** on a une taille n, et on souhaite tirer des entiers de longueur n (entre 2^n et 2^{n+1}), premiers, aléatoires uniformes.

Tirage d'entiers (presque) premiers aléatoires

- **Le problème :** on a une taille n, et on souhaite tirer des entiers de longueur n (entre 2^n et 2^{n+1}), premiers, aléatoires uniformes.
- D'après le théorème des nombres premiers, l'algorithme consistant à prendre des entiers de longueur n aléatoires, et de les rejeter tant qu'ils ne sont pas premiers, est raisonnable : il devrait faire de l'ordre de n rejets en moyenne.

Tirage d'entiers (presque) premiers aléatoires

- **Le problème :** on a une taille n, et on souhaite tirer des entiers de longueur n (entre 2^n et 2^{n+1}), premiers, aléatoires uniformes.
- D'après le théorème des nombres premiers, l'algorithme consistant à prendre des entiers de longueur n aléatoires, et de les rejeter tant qu'ils ne sont pas premiers, est raisonnable : il devrait faire de l'ordre de n rejets en moyenne.
- Que se passe-t-il si à la place d'un "vrai" test exact de primalité, on met un algorithme Monte Carlo comme le test de Miller-Rabin ?

Tirage d'entiers (presque) premiers aléatoires

- **Le problème :** on a une taille n, et on souhaite tirer des entiers de longueur n (entre 2^n et 2^{n+1}), premiers, aléatoires uniformes.
- D'après le théorème des nombres premiers, l'algorithme consistant à prendre des entiers de longueur n aléatoires, et de les rejeter tant qu'ils ne sont pas premiers, est raisonnable : il devrait faire de l'ordre de n rejets en moyenne.
- Que se passe-t-il si à la place d'un "vrai" test exact de primalité, on met un algorithme Monte Carlo comme le test de Miller-Rabin ?
  - Les nombres premiers ont toujours la même probabilité d'apparaître

Tirage d'entiers (presque) premiers aléatoires

- **Le problème :** on a une taille n, et on souhaite tirer des entiers de longueur n (entre 2^n et 2^{n+1}), premiers, aléatoires uniformes.
- D'après le théorème des nombres premiers, l'algorithme consistant à prendre des entiers de longueur n aléatoires, et de les rejeter tant qu'ils ne sont pas premiers, est raisonnable : il devrait faire de l'ordre de n rejets en moyenne.
- Que se passe-t-il si à la place d'un "vrai" test exact de primalité, on met un algorithme Monte Carlo comme le test de Miller-Rabin ?
  - Les nombres premiers ont toujours la même probabilité d'apparaître
  - On prend le risque de voir apparaître des non premiers

Tirage d'entiers (presque) premiers aléatoires

- **Le problème :** on a une taille n, et on souhaite tirer des entiers de longueur n (entre 2^n et 2^{n+1}), premiers, aléatoires uniformes.
- D'après le théorème des nombres premiers, l'algorithme consistant à prendre des entiers de longueur n aléatoires, et de les rejeter tant qu'ils ne sont pas premiers, est raisonnable : il devrait faire de l'ordre de n rejets en moyenne.
- Que se passe-t-il si à la place d'un "vrai" test exact de primalité, on met un algorithme Monte Carlo comme le test de Miller-Rabin ?
  - Les nombres premiers ont toujours la même probabilité d'apparaître
  - On prend le risque de voir apparaître des non premiers
  - Il convient de *calibrer* la probabilité d'erreur du test (le nombre de répétitions k) pour garantir une faible probabilité que ce soit un nombre non premier qui sorte

Tirage d'entiers (presque) premiers aléatoires

- **Le problème :** on a une taille n, et on souhaite tirer des entiers de longueur n (entre 2^n et 2^{n+1}), premiers, aléatoires uniformes.
- D'après le théorème des nombres premiers, l'algorithme consistant à prendre des entiers de longueur n aléatoires, et de les rejeter tant qu'ils ne sont pas premiers, est raisonnable : il devrait faire de l'ordre de n rejets en moyenne.
- Que se passe-t-il si à la place d'un "vrai" test exact de primalité, on met un algorithme Monte Carlo comme le test de Miller-Rabin ?
  - Les nombres premiers ont toujours la même probabilité d'apparaître
  - On prend le risque de voir apparaître des non premiers
  - Il convient de *calibrer* la probabilité d'erreur du test (le nombre de répétitions k) pour garantir une faible probabilité que ce soit un nombre non premier qui sorte
- Autre solution : ajouter un test exact, mais plus coûteux, **après** le test probabiliste - avec probabilité proche de 1, ce test coûteux ne sera exécuté qu'une seule fois.