

Rapport

Sonny Klotz - Idir Hamad - Younes Benyamna - Malek Zemni

Projet M1 Informatique
Primalité

29/04/2018



Module *TER*

Table des matières

1	Architecture de l'application	1
1.1	Organigramme et données échangées	1
1.2	Fonctionnalités des modules	2
1.3	Outils et langages de programmation	3
2	Cryptosystèmes - RSA	3
2.1	Description de RSA	4
2.2	Rôle des nombres premiers	4
3	Génération des nombres premiers	5
4	Tests de primalité	6
4.1	Test naïf - Crible d'Eratosthène	7
4.2	Test de Wilson	8
4.3	Test de Fermat	9
4.3.1	Algorithme	9
4.3.2	Complexité	11
4.3.3	Preuve	11
4.4	Test de Miller-Rabin	11
4.4.1	Algorithme	11
4.4.2	Complexité	13
4.4.3	Preuve	13
4.5	Test de Solovay-Strassen	13
4.5.1	Algorithme	13
4.5.2	Complexité	15
4.5.3	Preuve	15
4.6	Test AKS	15
4.6.1	Algorithme	15
4.6.2	Complexité	15
4.6.3	Preuve	15
5	Mesures de performance et comparatifs	15
5.1	Évolution des tests de primalité	16
5.2	Mesure du temps d'exécution	16
5.3	Génération optimale d'un nombre premier	19

Table des figures

1	Organigramme des différents modules de l'application	2
2	Processus de génération des nombres premiers	5
3	La fonction $\pi(n)$ pour les 1000 premiers nombres premiers	5
4	Comparatif des courbes représentatives du temps d'exécution de chaque test de primalité en fonction de la taille en bits du nombre premier à générer	18

Liste des Algorithmes

1	Test naïf	7
2	Crible d'Eratosthène	8
3	Test de Wilson	9
4	Test de Fermat	10
5	Test de Miller-Rabin	12
6	Test de Solovay-Strassen	14
7	Mesure temps exécution	17
8	RPNG Optimal	19

Liste des théorèmes et des définitions

1	Théorème (Théorème des Nombres Premiers)	5
2	Théorème (Théorème de Wilson)	8
3	Théorème (Petit théorème de Fermat (énoncé 1))	9
4	Théorème (Petit théorème de Fermat (énoncé 2))	9
1	Définition (Témoin de Fermat)	10
2	Définition (Nombre pseudo-premier)	10
3	Définition (Nombre de Carmichael)	10
5	Théorème (Critère d'Euler)	13
4	Définition (Résidu quadratique)	13
5	Définition (Symbole de Legendre)	14
6	Définition (Symbole de Jacobi)	14
7	Définition (Témoin d'Euler)	14
8	Définition (Nombre pseudo-premier d'Euler-Jacobi)	15

Introduction

Ce document est le compte-rendu final de notre projet sur les tests de primalité qui s'inscrit dans le cadre du module *TER* du M1 informatique de l'*UVSQ*.

Les tests de primalité sont des algorithmes qui permettent de savoir si un nombre entier est premier. Ces tests sont indispensables pour la cryptographie à clé publique.

Il existe plusieurs algorithmes de tests de primalité. L'efficacité de ces algorithmes est particulièrement liée à la taille des données fournies.

Notre travail consiste donc à implémenter différents tests de primalité et de comparer leurs performances. Ces mesures de performance nous permettront de construire un test de primalité optimal, produit final de ce projet.

Dans la première partie de ce document, on présentera l'architecture de notre application, illustrée par un organigramme.

Ensuite, on parlera des principaux cryptosystèmes faisant appel à des tests de primalité.

La troisième partie traitera des différents algorithmes de tests de primalité implémentés.

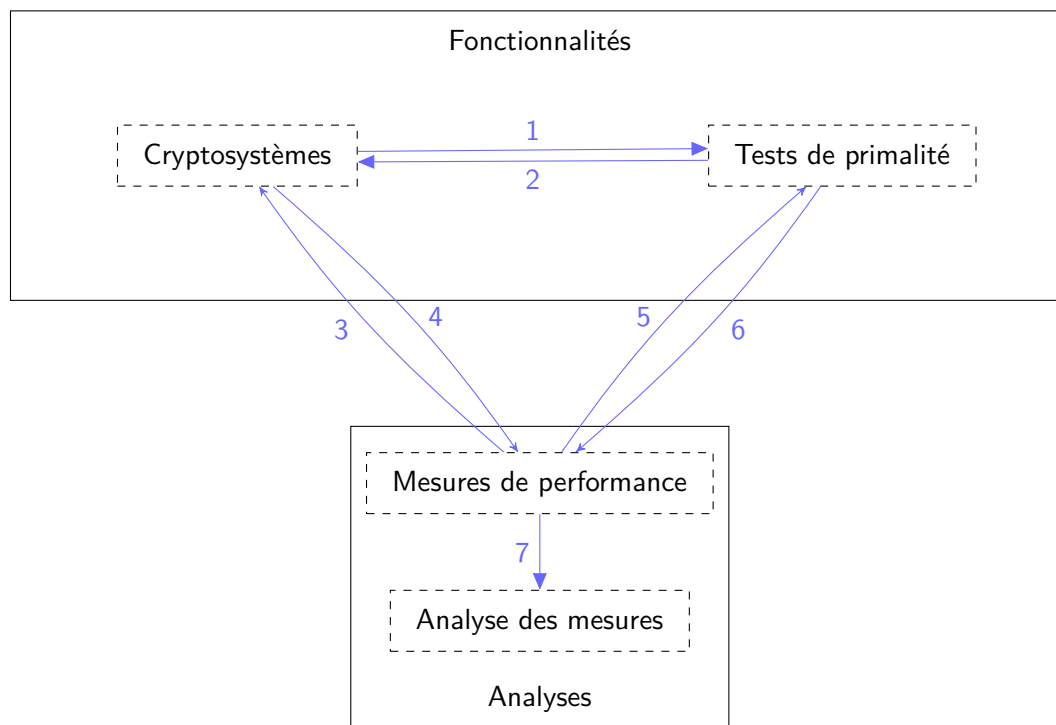
Les mesures de performance et le comparatif des tests de primalités seront détaillés dans la quatrième partie.

Finalement, on établira un bilan technique de notre projet, quant à l'application, à l'organisation interne au sein du groupe et aux coûts.

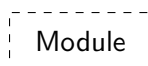
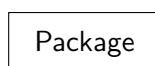
1 Architecture de l'application

1.1 Organigramme et données échangées

Cet organigramme représente la décomposition en modules de l'application ainsi que les informations qui circulent entre ces modules.



Légende :



informations transmises

FIGURE 1 – Organigramme des différents modules de l'application

Notes :

- (1) Nombre entier à tester (primalité)
- (2) Réponse sur la primalité (0 composé, 1 probablement premier, 2 premier)
- (3) Message à chiffrer
- (4) Chiffré du message
- (5) Nombre entier à tester (primalité)
- (6) Réponse sur la primalité (0 composé, 1 probablement premier, 2 premier)
- (7) Données collectées des différentes mesures de performance

1.2 Fonctionnalités des modules

Package Fonctionnalités

1. Module Cryptosystèmes : implémentation de cryptosystèmes ayant recours à des nombres premiers (**RSA**) et des générateurs de nombres premiers.
2. Module Test de primalité : implémentation de différents algorithmes de tests de primalité qui feront l'objet d'une étude comparative par la suite :
 - Test naïf

- Test de Wilson
- Test de Fermat
- Test de Miller-Rabin
- Test de Solovay-Strassen
- Test AKS

Package Analyses

1. Module Mesures de performance : mesures des performances des différents tests de primalité implémentés selon les valeurs données en entrées et les cryptosystèmes qui les utilisent.
2. Module Analyse des mesures :
 - Analyse comparative des différentes mesures calculées par le module Mesures de performance.
 - Produit final de l'application.

1.3 Outils et langages de programmation

Notre application va être implémentée dans le langage C. Le langage C possède plusieurs types pour représenter des nombre entiers. Cependant, tous ces types ont une précision fixe et ne peuvent pas dépasser un certain nombre d'octets. Le type le plus grand est le `long long int` qui peut contenir des entiers d'un taille maximale de 64 bits. Or, tous ces types sont beaucoup trop courts pour les applications cryptographiques qui nécessitent la manipulation de données d'au moins 512 bits.

GNU MP pour GNU Multi Precision, souvent appelée GMP est une bibliothèque C/C++ de calcul multiprécision sur des nombres entiers, rationnels et à virgule flottante qui permet en particulier de manipuler de très grand nombres.

Finalement, le logiciel Gnuplot va être utilisé pour faire des représentations graphiques à partir des résultats issus de notre programme.

2 Cryptosystèmes - RSA

Les tests de primalité sont des algorithmes indispensables pour la cryptographie à clé publique. Ces tests sont couramment utilisés par les cryptosystèmes **RSA** et **ElGamal** afin de générer des nombres premiers.

Pour **RSA**, les tests sont effectués lors la phase de génération de clés. Pour **ElGamal**, ils sont effectués lors de l'établissement d'un échange de clés.

Dans cette partie, on va détailler le cryptosystème **RSA** et exhiber rôle important des nombres premiers. On a choisi de s'intéresser qu'à un seul cryptosystème puisque le choix du cryptosystème n'aura pas d'effet sur les performances des tests de primalité, objet principal de ce projet.

2.1 Description de RSA

Décrit en 1977 par Ronald Rivest, Adi Shamir et Leonard Adleman, RSA est un cryptosystème basé sur le problème de factorisation, qui utilise une paire de clés (publique, privée) permettant de chiffrer et de déchiffrer un message. Le fonctionnement de RSA peut être décrit en 3 phases :

1. Génération des clés

- Choisir 2 grands **nombres premiers** distincts p et q .
- Calculer $n = p * q$. n est le module RSA et fait 1024 bits au minimum en général.
- Calculer $\Phi(n) = (p - 1)(q - 1)$.
- Choisir $e \in \mathbb{Z}_{\Phi(n)}^*$ (e premier avec $\Phi(n)$).
- Calculer d telle que $d * e \equiv 1 \text{ mod } \Phi(n)$ (d inverse de e pour la multiplication modulo $\Phi(n)$).

Les éléments échangés constituant la clé publique sont (n, e) . Les éléments constituant la clé privée sont (p, q, d) .

2. Chiffrement

Pour chiffrer un message M en un chiffré C , on utilise les éléments de la clé publique (n, e) :

$$C \equiv M^e \pmod{n}$$

3. Déchiffrement

Pour déchiffrer un chiffré C en un message clair M , on utilise les éléments de la clé privée (p, q, d) :

$$M \equiv C^d \pmod{n}$$

2.2 Rôle des nombres premiers

La première étape pour la mise en place d'un cryptosystème RSA est la génération de deux très grands nombres premiers p et q . Leur produit $n = p * q$ forme le module RSA. Pour cette raison, la taille de p et q en bits, doit être égale à la moitié de la taille en bits du module n . Par exemple, dans le cadre de RSA-1024, les deux nombres premiers doivent avoir une longueur de 512 bits.

En effet, un attaquant qui connaît le module RSA n et la clé publique e doit connaître la factorisation de n en nombres premiers pour trouver la clé privée d . Ainsi, l'entier n doit être très grand afin que sa factorisation ne soit pas possible avec les ressources de calcul actuelles. On voit donc l'intérêt crucial pour la sécurité de générer les deux grands nombres premiers p et q .

Parmi les algorithmes classiques de factorisation les plus efficaces, on retrouve **GNFS** (General Number Field Sieve) dont le temps d'exécution croît exponentiellement à la taille de n (complexité exponentielle). Avec les puissances de calcul actuelles, il est de plus en plus déconseillé d'utiliser un module RSA de taille 1024 bits. Il est estimé qu'un module de taille 2048 bits soit sécurisé (complexité factorisation supérieure à 2^{80}) jusqu'à l'année 2020. En 1994, l'algorithme de Shor appliqué sur des ordinateurs quantiques a permis d'effectuer une factorisation en un temps non exponentiel. Les applications des ordinateurs quantiques permettent théoriquement de casser RSA par la force brute, mais actuellement ces ordinateurs génèrent des erreurs aléatoires qui les rendent inefficaces.

3 Génération des nombres premiers

Les cryptosystèmes utilisent une approche commune pour la génération des nombres premiers. Cette approche générale consiste à utiliser un générateur de nombres aléatoires pour générer un entier, dont on testera ensuite la primalité. Ce processus est illustré par la figure ci-dessous :

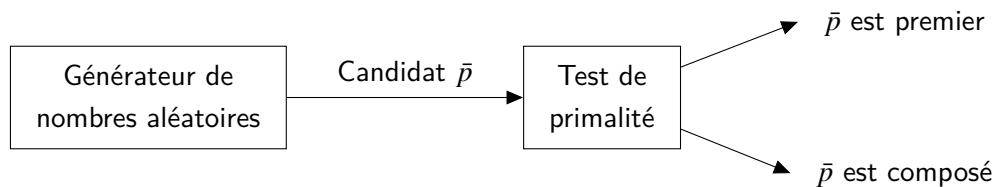


FIGURE 2 – Processus de génération des nombres premiers

Dans cette démarche, il est important d'utiliser un bon générateur de nombres aléatoires, qui ne doit dans aucun cas être prévisible. Si un attaquant réussit à deviner les nombres premiers qui composent le module RSA, alors le système est immédiatement cassé.

Fréquence des nombres premiers

Lors de la génération des nombres premiers à l'aide de ce processus, on voudrait savoir combien de nombres doit-on tester avant de trouver un nombre premier. La réponse à cette question est donnée par le *Théorème des Nombres Premiers*.

Théorème 1 (Théorème des Nombres Premiers). *soit $\pi(n)$ le nombre de premiers qui sont inférieurs à n , alors*

$$\pi(n) \approx \frac{n}{\ln(n)} \quad (n \rightarrow +\infty)$$

Un graphique de la fonction $\pi(n)$ pour les 1000 premiers nombres premiers est donné dans la figure ci-dessous :

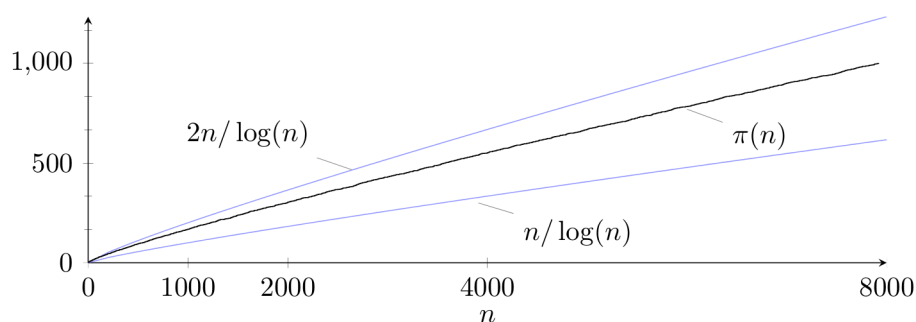


FIGURE 3 – La fonction $\pi(n)$ pour les 1000 premiers nombres premiers

Le tableau suivant contient l'approximation ainsi que le nombre exact de nombres premiers pour différentes valeurs de n . On y remarque que l'approximation est assez bonne.

n	$n/\ln(n)$	$\pi(n)$
10^3	145	168
10^4	1086	1229
10^5	8686	9592
10^6	72382	78498
10^7	620420	664579

Probabilité de générer un nombre premier p de k bits : on sait que $2^{k-1} \leq p \leq 2^k - 1$. Le nombre de nombres premiers dans cet intervalle (c-à-d de k bits) peut être approximé par :

$$\pi(2^k) - \pi(2^{k-1}) \approx \frac{2^k}{\ln(2^k)} - \frac{2^{k-1}}{\ln(2^{k-1})} \approx \frac{2^{k-1}}{\ln(2^{k-1})}$$

puisque $\ln(2^k) = \ln(2 \cdot 2^{k-1}) = \ln(2) + \ln(2^{k-1})$, donc $\ln(2^k) \approx \ln(2^{k-1})$ pour k grand.

Donc, il y'a 2^{k-1} entiers $\in [2^{k-1}, 2^k - 1]$ (de k bits) dont approximativement $\frac{2^{k-1}}{\ln(2^{k-1})}$ parmi eux qui sont premiers. Par conséquent, un nombre p de k bits sera premier avec une probabilité de :

$$\frac{1}{\ln(2^{k-1})}$$

Cas de RSA-1024 : pour générer une des deux clé de RSA-1024 dont la taille en bits est 512, on a probabilité de $1/\ln(2^{511}) \approx 1/355$ pour générer un nombre premier de 512 bits. Cette chance double si on se restreint sur les entiers impairs, c'est à dire qu'on doit générer à peu près 177 nombres avant de tomber sur un nombre premier.

4 Tests de primalité

Les tests de primalité interviennent dans la deuxième étape du processus de génération des nombres premiers. Ce sont des algorithmes qui permettent de savoir si un nombre entier est premier. Dans le cas où le nombre n'est pas premier, il est dit **composé**. Dans cette partie, on va détailler différents algorithmes de tests de primalité.

Les tests de primalité peuvent être :

- **déterministes** : fournissent toujours la même réponse pour un nombre donné
- **probabilistes** : peuvent fournir des réponses différentes pour un même nombre (utilisent des données tirées aléatoirement)

Voici la liste des différents algorithmes de tests de primalité qu'on va plus ou moins aborder :

Algorithme	Année	Type
Naïf (Crible d'Eratosthène)	-240	Déterministe
Fermat	1640	Probabiliste
Wilson	1770	Déterministe
Miller-Rabin	1976	Probabiliste
Solovay-Strassen	1977	Probabiliste
AKS	2002	Déterministe

Certains tests seront énoncés rapidement du fait qu'il ne sont pas assez performants. Par contre, on s'intéressera plus en détail aux tests de *Fermat*, *Miller-Rabin*, *Solovay-Strassen* et *AKS*. Pour chacun de ces tests, on donnera un bref historique, son algorithme, sa complexité, sa preuve, ainsi que son implémentation.

4.1 Test naïf - Crible d'Eratosthène

Le test naïf représente l'idée la plus intuitive pour tester la primalité d'un nombre entier. Pour décider si un nombre n est premier ou composé, on teste si les entiers $2, 3, \dots, n-1$ divisent n . Si un parmi ces entiers divise n alors on déduit que n est composé, sinon on conclut qu'il est premier. Ceci revient à factoriser le nombre en question.

Pour améliorer cet algorithme, on sait qu'un diviseur d'un entier n quelconque ne peut dépasser $n/2$. De plus, si n possède un diviseur plus grand que \sqrt{n} , alors il a forcément au moins un diviseur plus petit que \sqrt{n} . On peut donc accélérer la recherche en ne prenant en compte que des nombres premiers inférieurs à \sqrt{n} . Pour cela il suffit de pré-calculer et de stocker dans une table tous les nombres premiers $\leq \sqrt{n}$. Le **crible d'Eratosthène** par exemple peut être utilisé dans ce but.

Algorithme 1 : Test naïf

Données : un entier n

pour tout nombre premier $p \leq \sqrt{n}$ **faire**

si p divise n **alors**
 retourner composé;

retourner premier;

Crible d'Eratosthène

Ce crible est un procédé établi par Eratosthène, un mathématicien grec du III^e siècle av. J.-C., qui permet de trouver tous les nombres premiers inférieurs à un certain entier naturel donné N . Dans notre cas, cet entier donné est \sqrt{n} , n étant le nombre dont on va tester la primalité.

L'algorithme procède par élimination : il s'agit de supprimer d'une table des entiers de 2 à N tous les multiples d'un entier. En supprimant tous les multiples, à la fin il ne restera que les entiers qui ne sont multiples d'aucun entier, et qui sont donc les nombres premiers.

- retirer les multiples du plus petit entier premier restant (multiples de 2, puis de 3, etc.)
- on peut s'arrêter lorsque le carré de ce plus petit entier premier restant est supérieur au plus grand entier premier restant, car dans ce cas, tous les non-premiers ont déjà été retirés précédemment

- à la fin du processus, tous les entiers qui n'ont pas été rayés sont les nombres premiers inférieurs à N

L'algorithme du crible est le suivant :

Algorithme 2 : Crible d'Eratosthène

Données : un entier N qui correspond à \sqrt{n}

Créer une liste L de couples (*entier, primalité*), pour les entiers allant de 2 jusqu'à N , avec une primalité initialisée à "premier" : $L = \{(2, \text{premier}), (3, \text{premier}), \dots, (N, \text{premier})\}$;

plusGrandPremier = N ;

pour tout nombre p marqué "premier" de la liste L (de manière croissante) **faire**

si $p^2 > \text{plusGrandPremier}$ **alors**

retourner L ;

$i = 2$;

tant que $p * i < N$ **faire**

 Marquer "composé" l'entier à la position $p * i$;

 Mettre à jour **plusGrandPremier** ;

$i++$;

Complexité

La complexité en temps de l'algorithme 1 (Test naïf) dans le pire des cas est de $\pi(n) \approx \frac{2\sqrt{n}}{\ln(n)}$ division, c'est-à-dire $O(\sqrt{n})$ opérations. Dans le cas de RSA-1024, la complexité de cette méthode avoisine les 2^{503} divisions.

4.2 Test de Wilson

Le test de primalité de Wilson est un test déterministe basé sur le théorème suivant :

Théorème 2 (Théorème de Wilson). *Un entier $n > 1$ est un nombre premier si et seulement si*

$$(n-1)! + 1 \equiv 0 \pmod{n}$$

Ce théorème fournit une caractérisation des nombres premiers assez anecdotique et ne constitue pas un test de primalité efficace. Son principal intérêt réside dans son histoire et dans la relative simplicité de son énoncé et de ses preuves.

En effet, ce théorème était connu à partir du XVII^e siècle en Europe. En 1770, *John Wilson* redécouvre une conjecture de ce théorème et la publie. Ensuite, les mathématiciens *Lagrange*, *Euler* et *Gauss* le démontrent chacun à son tour.

Algorithme

L'algorithme du test de primalité basé sur le théorème de Wilson est le suivant :

Algorithme 3 : Test de Wilson

Données : un entier n

si $(n-1)! + 1 \equiv 0 \pmod{n}$ **alors**

retourner premier;

sinon

retourner composé;

Complexité

Ce test qui est basé sur une propriété très simple a cependant une complexité trop élevée. Il faut effectuer environ n multiplications modulaires, par conséquent la complexité est de $O(n)$.

4.3 Test de Fermat

Le test de Fermat est un test de primalité probabiliste basé sur le *petit théorème de Fermat* :

Théorème 3 (Petit théorème de Fermat (énoncé 1)). *Si p est un nombre premier, alors pour tout nombre entier a premier avec p*

$$a^{p-1} \equiv 1 \pmod{p}$$

Il existe un énoncé équivalent de ce théorème, qui est le suivant :

Théorème 4 (Petit théorème de Fermat (énoncé 2)). *Si p est un nombre premier, et a un nombre entier quelconque, alors*

$$a^p \equiv a \pmod{p}$$

Ce théorème doit son nom à *Pierre de Fermat*, qui l'énonce la première fois en 1640.

4.3.1 Algorithme

Le premier énoncé du théorème de Fermat va être exploité pour construire l'algorithme du test de primalité. Ce théorème décrit une propriété commune à tous les nombres premiers qui peut être utilisée pour détecter si un nombre est premier ou bien composé.

En effet, pour un entier n dont on veut tester la primalité et un entier a quelconque telle que $1 < a < n$:

- Le fait de choisir $1 < a < n$ garantit que si n était premier, a sera forcément premier avec n (puisque $a < n$) et ainsi le test n'échouera pas.
- Si $a^{n-1} \not\equiv 1 \pmod{n}$, alors n est surement composé.

Parmi les entiers a qui ne vérifient pas l'inégalité de Fermat, il y a évidemment ceux qui ne sont pas premiers avec n . Si l'on trouve un tel entier a (qu'il soit premier ou non avec n), on dit que a est un **témoin de non primalité** de n issu de la divisibilité (*témoin de Fermat*).

Définition 1 (Témoin de Fermat). Soit un entier $n \geq 2$. On appelle témoin de Fermat pour n , tout entier a , telle que

$$1 < a < n \quad \text{et} \quad a^{n-1} \not\equiv 1 \pmod{n}$$

- Si $a^{n-1} \equiv 1 \pmod{n}$, on ne peut pas conclure avec certitude que n est premier puisque la réciproque du *théorème de Fermat* est fausse (théorème 3).

Un nombre n vérifiant cette équation peut être premier, mais aussi composé, dans ce cas n est dit **pseudo-premier** de base a ou menteur.

Définition 2 (Nombre pseudo-premier). Un nombre pseudo-premier est un nombre premier probable (un entier naturel qui partage une propriété commune à tous les nombres premiers) qui n'est en fait pas premier. Un nombre pseudo-premier provenant du théorème de Fermat est appelé nombre pseudo-premier de Fermat.

Si un nombre pseudo-premier n de base a est pseudo-premier pour toutes les valeurs de a qui sont premières avec n est appelé **nombre de Carmichael**.

Définition 3 (Nombre de Carmichael). Un entier positif composé n est appelé nombre de Carmichael si pour tout entier a premier avec n ,

$$a^{n-1} \equiv 1 \pmod{n}$$

L'entier $n = 561 = 3 \cdot 11 \cdot 17$ est le plus petit nombre de Carmichael puisque $a^{560} \equiv 1 \pmod{561}$ pour tout entier a premier avec 561. Les nombres de Carmichael sont très rares. Il existe par exemple seulement 246 683 nombres de Carmichael inférieurs à 10^{16} . Le nombre de premiers inférieurs à 10^{16} est quant à lui égal à 279 238 341 033 925. Donc la probabilité qu'un nombre premier inférieur à 10^{16} soit un nombre de Carmichael est plus petite que $1/10^9$.

Les nombres pseudo-premiers et les nombre de Carmichael sont relativement rares. On peut donc envisager d'adopter ce critère pour un test probabiliste de primalité, qui est le test de Fermat. En effet, va être être répété k fois, et à chaque itération, on effectue le test avec une base a différente. Plus le nombre de répétitions est grand, plus la probabilité que le résultat du test soit correct augmente.

Algorithme 4 : Test de Fermat

Données : un entier n et le nombre de répétitions k

pour $i = 1$ *jusqu'à* k **faire**

 Choisir aléatoirement a tel que $1 < a < n$;

si $a^{n-1} \not\equiv 1 \pmod{n}$ **alors**

retourner composé;

retourner probablement premier;

4.3.2 Complexité

Si un algorithme rapide est utilisé pour l'exponentiation modulaire (par exemple Square&Multiply), la complexité en temps de l'algorithme de Fermat est

$$O(k \cdot \log_2(n) \cdot \log(\log(n)) \cdot \log(\log(\log(n))))$$

où k est le nombre de fois que le test est répété.

4.3.3 Preuve

Sonny

4.4 Test de Miller-Rabin

Le test de Miller-Rabin est un autre test de primalité probabiliste basé sur le *petit théorème de Fermat* (théorème 3). Il exploite quant à lui quelques propriétés supplémentaires. La version originale de ce test, publiée par Gary L. Miller en 1976, est déterministe, mais ce déterminisme dépend d'une hypothèse non démontrée (hypothèse de Riemann généralisée). En 1980, Michael Rabin a modifié cette hypothèse pour obtenir un algorithme de test probabiliste inconditionnel.

4.4.1 Algorithme

Dans le cas du test de Miller-Rabin, la propriété en question est un raffinement du *petit théorème de Fermat* (théorème 3). De la même façon que le test de Fermat, le test de Miller-Rabin tire parti d'une propriété de l'entier n dont on va tester la primalité, qui dépend d'un entier auxiliaire, le témoin a , et qui est vraie dès que n est un nombre premier. Le principe de ce test probabiliste est donc de vérifier cette propriété pour suffisamment de témoins.

En effet, soit $n > 2$ un entier dont on va tester la primalité et a un entier quelconque telle que $1 < a < n$. Soient $s \in \mathbb{N}^*$ et $t \in \mathbb{N}$ **impair**, on peut écrire $n - 1 = 2^s t$ (s est le nombre maximum de fois que l'on peut mettre 2 en facteur dans $n - 1$).

Alors, dans le cas où n est premier, d'après le *petit théorème de Fermat* (théorème 3) on a :

$$a^{n-1} \equiv 1 \pmod{n}$$

$$\Leftrightarrow$$

$$a^{2^s t} \equiv 1 \pmod{n}$$

$$\Leftrightarrow$$

$$a^{2^s t} - 1 \equiv 0 \pmod{n}$$

$$\Leftrightarrow$$

$$a^{2^s t} - 1 = (a^{2^{s-1} t})^2 - 1 = (a^{2^{s-1} t} + 1)(a^{2^{s-1} t} - 1) \equiv 0 \pmod{n} \quad (\text{puisque } s \geq 1)$$

Si $s - 1 > 0$ alors le dernier terme est de nouveau une différence de carrés qui peut donc être factorisée. En continuant de la même manière, on obtient au final l'expression suivante :

$$(a^{2^{s-1} t} + 1)(a^{2^{s-2} t} + 1) \dots (a^t + 1)(a^t - 1) \equiv 0 \pmod{n} \quad (*)$$

On sait que pour un nombre premier p , si $ab \equiv 0 \pmod{p}$ alors $a \equiv 0 \pmod{p}$ ou $b \equiv 0 \pmod{p}$. Par conséquent, si n est premier, alors l'équation (*) est vraie si et seulement si un des termes de sa partie gauche est $0 \pmod{n}$. Autrement dit, si n est premier alors

$$a^{2^j t} \equiv -1 \pmod{n} \quad \text{pour au moins un } j \in \{0, 1, \dots, s-1\}$$

ou

$$a^t \equiv 1 \pmod{n}$$

Si pour un nombre entier n , une des équations ci-dessus est vérifiée, alors l'algorithme conclue que n est probablement premier et termine. Si aucune de ces équations n'est vérifiée, alors l'algorithme renvoie que n est composé avec certitude. On peut aussi apporter une amélioration à cette approche.

Si on trouve que

$$a^{2^j t} \equiv 1 \pmod{n} \quad \text{pour un } j \in \{0, 1, \dots, s-1\},$$

on peut directement conclure que n est composé et terminer l'exécution de l'algorithme. Ceci est dû au fait que pour un nombre p premier, les seuls éléments pour lesquels un entier x vérifie $x^2 \equiv 1 \pmod{p}$ sont 1 et -1 , qui sont les deux seules racines carrées de l'unité.

Compte tenu de tous les éléments développés ci-dessus, l'algorithme du test de primalité de Miller-Rabin est le suivant :

Algorithme 5 : Test de Miller-Rabin

Données : un entier n et le nombre de répétitions k

Décomposer $n-1 = 2^s t$, avec $s \in \mathbb{N}^*$ et $t \in \mathbb{N}$ impair ;

pour $i = 1$ *jusqu'à* k **faire**

 Choisir aléatoirement a tel que $1 < a < n$;

$y \leftarrow a^t \pmod{n}$;

si $y \not\equiv 1 \pmod{n}$ *et* $y \not\equiv -1 \pmod{n}$ **alors**

pour $j = 1$ *jusqu'à* $s-1$ **faire**

$y \leftarrow y^2 \pmod{n}$;

si $y \equiv 1 \pmod{n}$ **alors**

retourner composé;

si $y \equiv -1 \pmod{n}$ **alors**

 Arrêter la boucle de j et continuer avec le i suivant (sans renvoyer composé);

retourner composé;

retourner probablement premier;

Probabilité d'erreur et nombre d'itérations : si le test de Miller-Rabin renvoie composé, alors le nombre est effectivement composé. Il peut être démontré que si le test de Miller-Rabin dit que n est premier, le résultat est faux avec une probabilité inférieure à $1/4$. En effet, il existe des valeurs de a qui produiront de manière répétée des menteurs, qui indiqueront donc que n est premier alors qu'il est composé. On appelle un **témoin** fort pour n un entier a pour lequel

$$a^t \not\equiv 1 \pmod{n} \quad \text{et} \quad a^{2^j t} \not\equiv -1 \pmod{n} \quad \text{pour tout } j \in \{0, \dots, s-1\}$$

Il peut être montré qu'il existe toujours un témoin fort pour n'importe quel composé impair n , et qu'au moins $3/4$ de ces valeurs pour a sont des témoins forts pour la composition de n . Si on répète ce test k fois, la probabilité que le résultat soit toujours faux décroît très rapidement. La probabilité que le test renvoie premier à tort après k itérations est donc de $1/4^k$.

4.4.2 Complexité

Sonny

4.4.3 Preuve

Sonny

4.5 Test de Solovay-Strassen

Le test de Solovay-Strassen est un test de primalité probabiliste, publié par *Robert Solovay* et *Volker Strassen* en 1977. Ce test a une importance historique dans la démonstration de la faisabilité du cryptosystème RSA.

4.5.1 Algorithme

L'algorithme du test de Solovay-Strassen est essentiellement basé sur le **critère d'Euler**, un théorème qui énonce que :

Théorème 5 (Critère d'Euler). Soient $p > 2$ un nombre premier et a un entier premier avec p

— Si a est un résidu quadratique modulo p , alors $a^{\frac{n-1}{2}} \equiv 1 \pmod{p}$.

— Si a n'est pas un résidu quadratique modulo p , alors $a^{\frac{n-1}{2}} \equiv -1 \pmod{p}$.

Ceci se résume en utilisant le symbole de Legendre par :

$$a^{\frac{n-1}{2}} \equiv \left(\frac{a}{p}\right) \pmod{p}$$

Définition 4 (Résidu quadratique). On dit qu'un entier q est un résidu quadratique modulo p s'il existe un entier x tel que :

$$x^2 \equiv q \pmod{p}$$

Autrement dit, un résidu quadratique modulo p est un nombre qui possède une racine carrée de module p . Dans le cas contraire, on dit que q est un non-résidu quadratique modulo p .

Le **symbole de Legendre** est utilisé pour résumer le critère d'Euler. Il est défini de la manière suivante :

Définition 5 (Symbole de Legendre). Le symbole de Legendre est une fonction de deux variables entières à valeurs dans $\{-1, 0, 1\}$. Si p est un nombre premier et a un entier, alors le symbole de Legendre $\left(\frac{a}{p}\right)$ vaut :

$$\begin{cases} 0 & \text{si } a \text{ est divisible par } p \\ 1 & \text{si } a \text{ est un résidu quadratique modulo } p \text{ mais pas divisible par } p \\ -1 & \text{si } a \text{ n'est pas un résidu quadratique modulo } p \end{cases}$$

Le cas particulier $p = 2$ est inclus dans cette définition mais est sans intérêt : $\left(\frac{a}{p}\right)$ vaut 0 si a pair et 1 sinon.

Pour pouvoir exploiter le *critère d'Euler* dans l'algorithme du test de primalité, on doit pouvoir calculer le *symbole de Legendre* pour tout entier n dont on veut tester la primalité. On introduit donc le **symbole de Jacobi** qui est une généralisation du *symbole de Legendre*, définit de la manière suivante :

Définition 6 (Symbole de Jacobi). Le symbole de Jacobi $\left(\frac{a}{n}\right)$ est défini, $\forall a \in \mathbb{Z}$ et $n \in \mathbb{N}$ impair, comme produit de symboles de Legendre, en faisant intervenir la décomposition en facteurs premiers de n . Si $n = p_1 * p_2 * \dots * p_k$ pour $k \in \mathbb{N}$ telle que p_1, p_2, \dots, p_k sont des nombres premiers non nécessairement distincts, alors :

$$\left(\frac{a}{n}\right) = \left(\frac{a}{\prod_{i \in \{1, \dots, k\}} p_i}\right) = \prod_{i \in \{1, \dots, k\}} \left(\frac{a}{p_i}\right)$$

Compte tenu des notions décrites ci-dessus, on peut maintenant construire l'algorithme de test de primalité de Solovay-Strassen :

Algorithme 6 : Test de Solovay-Strassen

Données : un entier n impair et le nombre de répétitions k

pour $i = 1$ jusqu'à k **faire**

 Choisir aléatoirement a tel que $2 < a < n$;

$x \leftarrow \left(\frac{a}{n}\right)$;

si $x = 0$ ou $x \neq a^{\frac{n-1}{2}} \pmod{n}$ **alors**

retourner composé;

retourner probablement premier;

Cet algorithme exploite essentiellement le *critère d'Euler* (théorème 5). En effet, pour un entier n dont on veut tester la primalité et un entier a quelconque telle que $2 < a < n$:

— Si $a^{\frac{n-1}{2}} \not\equiv \left(\frac{a}{n}\right) \pmod{n}$, alors n est surement composé.

Parmi les entiers a qui ne vérifient pas le *critère d'Euler* (théorème 5), il y a évidemment ceux qui ne sont pas premiers avec n . Si l'on trouve un tel entier a (qu'il soit premier ou non avec n), on dit que a est un **témoin de non primalité** de n (*témoin d'Euler*).

Définition 7 (Témoin d'Euler). Soit un entier $n > 2$. On appelle *témoin d'Euler* pour n , tout entier a , telle que

$$2 < a < n \quad \text{et} \quad a^{\frac{n-1}{2}} \not\equiv \left(\frac{a}{n}\right) \pmod{n}$$

- Si $a^{\frac{n-1}{2}} \equiv \left(\frac{a}{n}\right) \pmod{n}$, on ne peut pas conclure avec certitude que n est premier puisque la réciproque du *critère d'Euler* (théorème 5) est fausse.

Un nombre n vérifiant cette équation peut être premier, mais aussi composé, dans ce cas n est dit **pseudo-premier d'Euler-Jacobi** de base a ou menteur.

Définition 8 (Nombre pseudo-premier d'Euler-Jacobi). *Un nombre pseudo-premier d'Euler-Jacobi de base a est un nombre composé impair n premier avec a et tel que la congruence suivante soit vérifiée :*

$$a^{\frac{n-1}{2}} \equiv \left(\frac{a}{n}\right) \pmod{n}$$

À la différence du test de primalité de Fermat, pour chaque entier composé n , au moins la moitié de tous les a sont des témoins d'Euler. Par conséquent, il n'y a aucune valeur de n pour laquelle tous les a sont des menteurs, alors que c'est le cas pour les nombres de *Carmichael* dans le test de Fermat.

4.5.2 Complexité

Sonny

4.5.3 Preuve

Sonny

4.6 Test AKS

4.6.1 Algorithme

Younes + Idir + Sonny

4.6.2 Complexité

Sonny

4.6.3 Preuve

Sonny

5 Mesures de performance et comparatifs

Dans la partie précédente, on s'est contenté de présenter différents tests de primalité sans conclure par rapport à leurs performances. Dans cette partie, on va mesurer ces performances et établir un comparatif entre les différents tests abordés.

5.1 Évolution des tests de primalité

Sonny développe cette partie quand tu fais les complexités

Plusieurs algorithmes de tests de primalité se sont succédés au fil du temps. Des algorithmes de plus en plus efficaces apparaissent pour en remplacer d'autres.

Premiers tests déterministes

Les premiers algorithmes de tests de primalité apparus sont les algorithmes de **test naïf**. Ces algorithmes sont déterministes, c'est-à-dire qu'ils retournent toujours une réponse exacte. Ils constituent la façon la plus naturelle de tester la primalité d'un nombre. Cependant, leur complexité était trop élevée, surtout quand il s'agissait de tester de grands nombres.

D'autres algorithmes déterministes très simples et un peu plus performant sont aussi apparus, telle que le **test de Wilson**. Mais la complexité est restée trop élevée surtout pour de grands nombres.

Tests probabilistes

Les algorithmes de tests probabilistes constituent une façon beaucoup plus efficace que les premiers algorithmes déterministes découverts pour tester la primalité d'un nombre. Ces algorithmes de type *Monte-Carlo* décident si un entier est premier ou pas avec une certaine probabilité P . La sortie d'un test de primalité probabiliste sur un entier n est soit :

- n est composé : toujours vrai
- n est premier : vrai avec une probabilité P

Dans le cas où la sortie du test est "premier", il y a toujours une petite probabilité que le nombre testé ne soit pas vraiment premier. Pour pallier à ce problème et diminuer cette probabilité, on a tendance à répéter le test probabiliste un certain nombre de fois.

Le premier algorithme probabiliste apparu est le **test de Fermat**. Ce test qui repose sur un théorème énoncé en 1640 a longtemps été utilisé en pratique, jusqu'à l'apparition du **test de Solovay-Strassen** en 1977.

À partir de 1980, le **test de Solovay-Strassen** a été lui aussi remplacé en pratique par le **test de Miller-Rabin**, plus efficace, car reposant sur un critère analogue, mais ne donnant de faux positif qu'au plus une fois sur quatre lorsque le nombre testé n'est pas premier.

Test déterministe rapide : AKS

Sonny

5.2 Mesure du temps d'exécution

Pour comparer les performances des algorithmes de test de primalité implémentés, on va mesurer le temps d'exécution de chaque algorithme, c'est-à-dire le temps de réponse, selon le nombre de bits de l'entier testé. Cette méthode va donc permettre de connaître le test le plus performant pour un nombre de bits donné.

Algorithme de mesure

L'algorithme qui correspond à cette méthode de mesure du temps d'exécution est le suivant :

Algorithme 7 : Mesure temps exécution

Données : tableau `tps[6][1025]` à remplir, correspondant au temps d'exécution des 6 test pour un nombre de bits entre 0 et 1024

Sorties : tableau `tps[6][1025]` rempli

pour *chaque test* ($i = 0$ jusqu'à 5) **faire**

pour *pour un nombre de bits allant de 0 à 1024* ($j = 0$ jusqu'à 1024) **faire**

 Générer un nombre premier p de j bits;
 `tps[i][j]` \leftarrow temps pour que le test i vérifie p ;

Cet algorithme est implémenté dans le fichier source [mesureTemps.c](#).

Analyse des mesures

Une fois que les mesures ont été effectuées, on va visualiser les résultats grâce à des représentations graphiques afin de les comparer et de conclure par rapport aux performances.

Le logiciel Gnuplot va être utilisé pour faire ces représentations graphiques. Pour cela, il faut écrire les résultats dans un fichier bien formaté dont voici un extrait :

```
1 # Test Fermat :
2 # BITS TEMPS
3 0 0.000000
4 1 0.000000
5 2 0.000001
6 3 0.000433
7 4 0.000483
8 5 0.000558
9 6 0.000573
10 7 0.000002
11 8 0.002930
12 9 0.015577
13 10 0.000989
14 11 0.001098
15 12 0.000739
16 13 0.000791
17 14 0.001213
18 15 0.001019
19 16 0.001044
20 17 0.000670
21 18 0.000475
22 19 0.000486
23 20 0.000657
24 21 0.001443
25 22 0.000475
26 23 0.000521
```

```

27 24 0.000617
28 25 0.000436
29 26 0.000421
30 27 0.000496
31 28 0.000579
32 29 0.000822
33 30 0.000666
34 31 0.000707
35 32 0.001029

```

Listing 1 – mesures.txt

Ce fichier contient 6 blocs de 1024 lignes, dont chaque bloc correspond à un test de primalité et dont les lignes correspondent au temps d'exécution mesuré pour le nombre de bit donné. En exploitant les résultat de ce fichier, le logiciel `Gnuplot` va permettre de construire les courbes représentatives du temps d'exécution de chaque test de primalité en fonction de la taille en bits du nombre premier à générer :

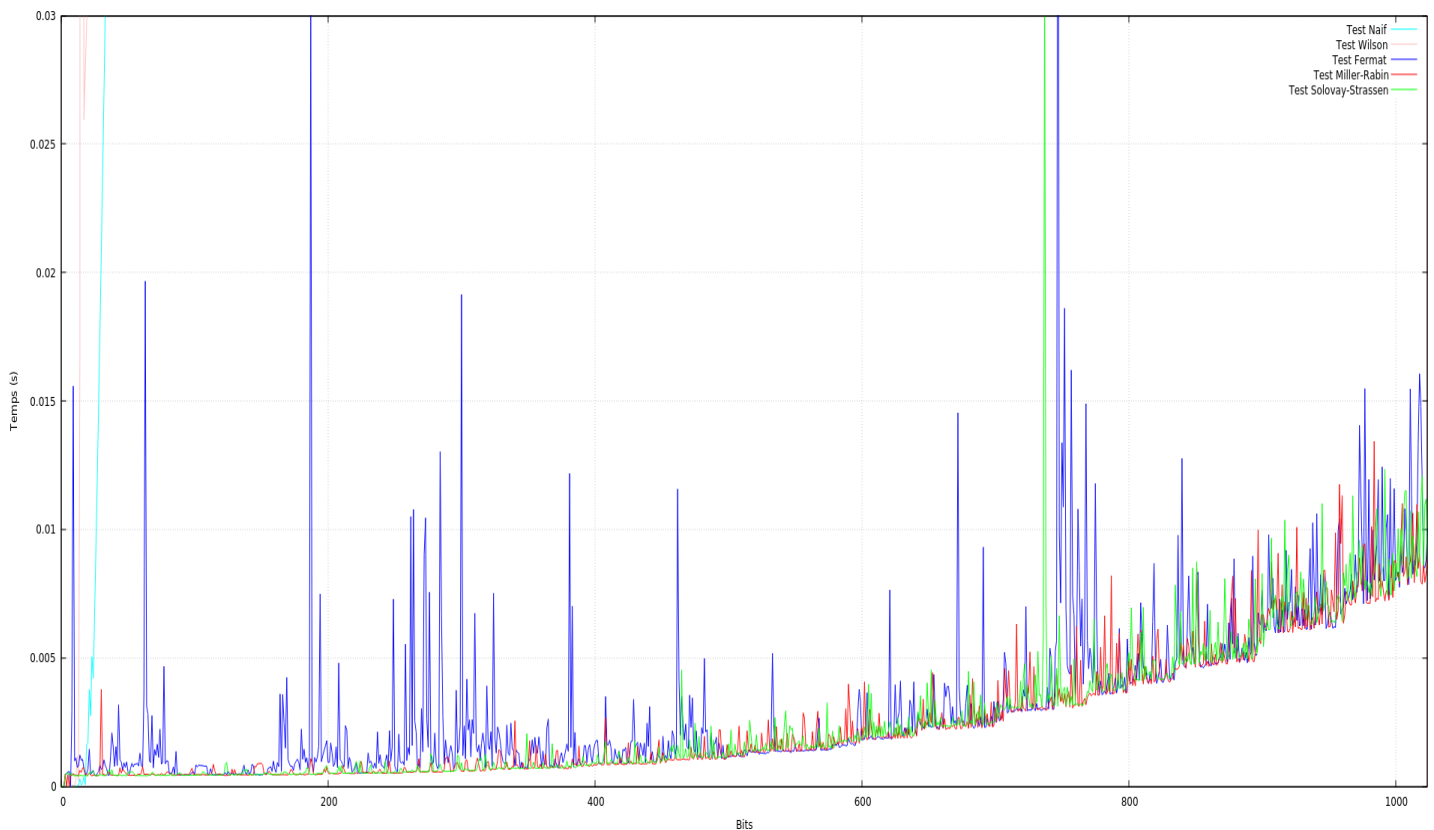


FIGURE 4 – Comparatif des courbes représentatives du temps d'exécution de chaque test de primalité en fonction de la taille en bits du nombre premier à générer

Ce graphique, qui est la superposition des courbes du temps d'exécution des 6 tests de primalité implémentés, va permettre d'observer les performances des différents algorithmes de test.

En effet, en ce qui concerne le **test Naif** et le **test de Wilson**, ceux-ci ont un temps de réponse rapide pour un nombre de bits petit, mais ce temps de réponse va croître d'une manière exponentielle

à partir de 16 bits, ce qui en fait des tests très lents globalement.

Concernant les tests probabilistes, ceux-ci ont un temps d'exécution rapide en moyenne. Le **test de Fermat** présente un temps assez variable qui s'avère très lent pour certain nombre de bits (on remarque la présence de nombreux pics qui sont parfois très hauts). Le **test de Solovay-Strassen** présente un temps variable lui aussi mais les pics sont d'une amplitude beaucoup moins importante par rapport au **test de Fermat**. Quant au **test de Miller-Rabin**, celui-ci inscrit un temps d'exécution faible en moyenne puisqu'il présente des pics avec de faibles amplitudes. De plus, il est le plus performant quand le nombre de bits est très grand.

5.3 Génération optimale d'un nombre premier

Compte tenu des résultats des mesures effectuées sur le temps d'exécution de chaque test de primalité, il est maintenant possible, lors de la génération d'un nombre premier de t bits, de choisir l'algorithme de test de primalité qui présente le temps d'exécution minimal pour t bits. Pour cela, on aura juste à exploiter les résultats pré-calculés dans la structure `tps[6][1025]` de la manière suivante :

Algorithme 8 : RPNG Optimal

Données : la taille t en bits de l'entier premier à générer et le tableau `tps[6][1025]` des résultats de mesure du temps d'exécution des 6 test pour un nombre de bits entre 0 et 1024

Sorties : un nombre premier de t bits

Trouver le test de primalité d'indice i (i entre 0 et 5) telle que `tps[i][t]` est minimale (c'est-à-dire le test le plus rapide pour vérifier t bits);

Générer un nombre premier de t bits avec le test de primalité d'indice i ;

Conclusion