Publicaciones Matemáticas del Uruguay
Volumen 15, Junio 2016, Páginas 123–139
S 0797-1443

# CÓDIGOS Y CRIPTOGRAFÍA: LA TEORÍA DE NÚMEROS APLICADA A TRES VIÑETAS DE AMOR

NATHAN C. RYAN

Resumen. La criptografía y la teoría de códigos son dos áreas de la teoría de números que tienen mucha aplicación a la comunicación. Describimos unos ejemplos destacados de las áreas nombradas, ejemplos que dan una vista panorámica de que ellas consisten.

## 1. Un cuento de amor y de códigos...

Nuestro héroe tímido Miguel toma el ómnibus. Una chica, desconocida para Miguel toma el mismo ómnibus. Desde el otro lado del bus Miguel la ve y piensa que es la chica más bonita que jamás haya visto. Ella está leyendo un libro, completamente absorta. Siendo nerd, además de ser tímido, Miguel logra memorizar el número ISBN del libro y anota

$$849838285X$$

después de que la chica se baja del ómnibus.

Miguel arma un plan: va a comprar el mismo libro y usará esta cosa en común para iniciar una conversación la próxima vez que ve a la chica en el ómnibus.

Pero...

Cuando llega a casa, entra el número de ISBN que escribió en un buscador de ISBN y lee

```
‘‘El número de ISBN ingresado tiene una cantidad incorrecta de
        dígitos o la suma verificadora equivocada.’’
```

**¿Qué es un número de ISBN?.** Un *International Standard Book Number* (ISBN) es una cadena de diez símbolos que únicamente identifica un libro. Los primeros nueve símbolos son dígitos y el décimo, un 'símbolo verificador' viene del conjunto

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, X\}.$$

¿Qué hace este símbolo verificador?

*Divisibilidad.* Arrancamos con un hecho que el lector ya debería saber del liceo, enunciado y nombrado como un teorema:

**Teorema** (El teorema cociente-resto)**.** *Sea* $a$ *un entero y* $x$ *un entero positivo. Existen enteros* $q$ *y* $r$ *únicos tal que*

1. $a = qx + r$*;*
2. $0 \leq r < x$.

Decimos que $q$ es el *cociente* y $r$ el *resto* cuando se divide $a$ por $x$.

Sean $x_1, x_2, \ldots, x_9$ los primeros nueve símbolo en el número ISBN. Sea $c$ el resto cuando la suma

$$\sum_{k=1}^{9} kx_k = 1 \cdot x_1 + 2 \cdot x_2 + 3 \cdot x_3 + 4 \cdot x_4 + 5 \cdot x_5 + 6 \cdot x_6 + 7 \cdot x_7 + 8 \cdot x_8 + 9 \cdot x_9$$

es dividida por 11. Como $c$ es el resto que resulta de una división por 11, $c$ es un elemento del conjunto

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}.$$

El símbolo verificador $x_{10}$ del ISBN es determinado por la regla:

$$x_{10} = \begin{cases} c & \text{si } 0 \leq c \leq 9 \\ X & \text{si } c = 10. \end{cases}$$

Se usa el símbolo X como lo usaba los romanos.

Ahora se nota que la suma $-c + \sum_{k=1}^{9} kx_k$ es divisible por 11.

*Máximo común divisor y mínimo común múltiplo.* El *maximal común divisor* de dos enteros $a$ y $b$, denotado $\mathrm{mcd}(a, b)$ es un entero $g$ que cumple

- $g \mid a$ y $g \mid b$;
- si $d \mid a$ y $d \mid b$, también $d \mid g$.

El *mínimo común múltiplo* de dos enteros $a$ y $b$, denotado $\mathrm{mcm}(a, b)$ es un entero $\ell$ que cumple

- $\ell$ es un múltiplo de $a$ y $\ell$ es un múltiplo de $b$;
- si $x$ es un múltiplo de $a$ y $x$ es un múltiplo de $b$, también $x$ es un múltiplo de $\ell$.

Se puede probar que

$$a \cdot b = \mathrm{mcm}(a, b) \cdot \mathrm{mcd}(a, b).$$

*Regresando al cuento de amor...* Se puede ver que el número que anotó Miguel es incorrecto porque la suma

$$\sum_{k=1}^{9} kx_k = 1 \cdot \mathbf{8} + 2 \cdot \mathbf{4} + 3 \cdot \mathbf{9} + 4 \cdot \mathbf{8} + 5 \cdot \mathbf{3} + 6 \cdot \mathbf{8} + 7 \cdot \mathbf{2} + 8 \cdot \mathbf{8} + 9 \cdot \mathbf{5} = 261$$

que da un resto de 8 cuando se divide por 11, hecho que implica el símbolo verificador debería ser un $8$ y no un X o que haya un error en los primeros nueve dígitos.

Si Miguel no tuviera mala suerte, no tendría ningún tipo de suerte...

**Detectando errores comunes.** El número de ISBN consiste de 10 símbolos, pero con los primeros nueve se puede únicamente identificar el libro. El décimo símbolo, el símbolo verificador es redundante. ¿Por qué se usan más símbolos de los que se precisa? Debería haber buena razón para justificar lo extra.

Los dos errores más comunes en manejando números de ISBN son un símbolo alterado y la transposición de dos símbolos adyacentes. Por ejemplo, quiero escribir

$$556176988\mathrm{X},$$

pero escribo

$$556\mathbf{2}76988\mathrm{X} \quad \text{o} \quad 556\mathbf{71}6988\mathrm{X}.$$

En cada caso el símbolo verificador está mal considerando los primeros nueve símbolos, sugiriendo que hay un error en los primeros nueve símbolos.

Hay errores que no se puede detectar con este método. Por ejemplo, si quiero escribir

$$673329141X,$$

pero escribo

$$672329341X,$$

el símbolo verificador no me dice nada porque en cada caso el 'X' sí es el símbolo correcto.

**Afirmación.** *El número ISBN puede detectar un símbolo incorrecto.*

Es decir, sean $x_1, x_2, \ldots, x_9$ y $y_1, y_2, \ldots, y_9$ los primeros nueve símbolos de dos candidatos para un ISBN. Y sea $i$ un entero tal que

1. $1 \leq i \leq 9$;
2. $x_i \neq y_i$;
3. $x_k = y_k$ para $k \neq i$.

Entonces

$$\sum_{k=1}^{9} kx_k \text{ y } \sum_{k=1}^{9} ky_k$$

tienen restos diferentes cuando se dividen por 11.

Entonces, si, en vez de escribir

$$x_1 \ldots x_{i-1} x_i x_{i+1} \ldots x_{10},$$

escribo

$$x_1 \ldots x_{i-1} y_i x_{i+1} \ldots x_{10},$$

voy a saber que algo está mal porque $x_{10}$ no es el símbolo verificador correcto para

$$x_1 \ldots x_{i-1} y_i x_{i+1} \ldots x_{10}.$$

Vamos a demostrar esta afirmación. Pero primero, tenemos que hablar de divisibilidad y la aritmética modular.

*Divisibilidad.*

**Definición** ($a$ divide $b$)**.** Sea $a$ un entero distinto de cero y sea $b$ cualquier entero. Decimos que $a$ *divide* $b$, y escribimos $a \mid b$, si existe un entero $m$ tal que $b = am$; si no, decimos que $a$ *no divide* a $b$ y escribimos $a \nmid b$.

**Ejemplo.** Entonces $a \mid b$ si y solo si 0 es el resto cuando se divide $b$ por $a$. Por ejemplo $7 \mid 35$ porque $35 = 7 \cdot 5$; $7 \mid (-14)$ porque $-14 = 7 \cdot (-2)$; $7 \nmid 11$ porque $11 = 7 \cdot 1 + 4$.

*Regresando al cuento de amor...* Quizás, no se ha perdido toda la esperanza...

*Propiedades de la divisibilidad.* Un *primo* es un entero $p$ divisible por exactamente dos enteros. Los primos son como los átomos de los enteros; esta idea se encapsula en:

**Proposición.** *Sean $a$ y $b$ enteros y sea $p$ un primo. Si $p \mid ab$, entonces $p \mid a$ o $p \mid b$.*

*Demostración.* Si $a$ o $b$ es cero, entonces $p$ trivialmente divide el producto. Supongamos que $p \nmid a$ y vamos a probar que $p \mid b$. Sea $\ell = \text{mcm}(p, a)$. Como $\text{mcd}(p, a) = 1$, $\ell = |pa|$. Como $p \mid ab$ y $a \mid ab$, vemos que $ab$ es un múltiplo común de $p$ y $a$. Entonces

$$(\ell \mid ab) \to (pa \mid ab) \to (p \mid b).$$

□

Otra propiedad útil de los primos con respeto a la divisibilidad:

**Proposición.** *Sean $a$ y $b$ enteros y sea $x$ un entero positivo. Entonces, $a$ y $b$ tienen el mismo resto con dividir por $x$ si y sólo si $x \mid (a - b)$.*

*Demostración.* Probamos la parte directa primero: usando el teorema cociente-resto, tenemos:

$$a = q_1 x + r_1 \text{ y } b = q_2 x + r_2.$$

Como $r_1 = r_2$, tenemos:

$$a - b = q_1 x - q_2 x = x(q_1 - q_2)$$

y con esto concluimos que $x \mid (a - b)$.

Recíprocamente, usando el teorema cociente-resto, $a = q_1 x + r_1$ y $b = q_2 x + r_2$ donde $r_1 \neq r_2$. Entonces $(a - b) = (q_1 - q_2)x + r_1 - r_2$. Como $r_1 \neq r_2$, vemos que $x \nmid (a - b)$. □

*Detectando errores.*

**Afirmación** (De vuelta)**.** *El código ISBN detecta una dígito incorrecto.*

Sean $x_1, x_2, \ldots, x_9$ y $y_1, y_2, \ldots, y_9$ los primeros nueve símbolo y sea $i$ un entero tal que

1. $1 \leq i \leq 9$;
2. $x_i \neq y_i$;
3. $x_k = y_k$ para $k \neq i$.

Entonces

$$\sum_{k=1}^{9} k x_k \text{ y } \sum_{k=1}^{9} k y_k$$

tienen restos diferentes cuando se dividen por 11. Esta condición es equivalente a

$$11 \nmid \left( \sum_{k=1}^{9} k x_k - \sum_{k=1}^{9} k y_k \right).$$

¿Por qué nos ayuda esta observación? Primero introducimos una notación útil: escribimos $a \equiv b \pmod{n}$ (se lea $a$ y $b$ son iguales módulo $n$) si $n \mid b - a$ o,

equivalentemente, que $a$ y $b$, después de dividir por $n$ tienen el mismo resto. Se nota que $a \equiv 0 \pmod{n}$ es equivalente a $n \mid a$. La relación

$$11 \nmid \left( \sum_{k=1}^{9} k x_k - \sum_{k=1}^{9} k y_k \right)$$

se puede escribir como

$$\left( \sum_{k=1}^{9} k x_k - \sum_{k=1}^{9} k y_k \right) \not\equiv 0 \pmod{11}.$$

Se puede hacer aritmética módulo $n$: si $a \equiv b \pmod{n}$ y $c \equiv d \pmod{n}$, sabemos

- $a + c \equiv b + d \pmod{n}$;
- $a - c \equiv b - d \pmod{n}$; y
- $a \cdot c \equiv b \cdot d \pmod{n}$.

División es un poco más complicado porque, por ejemplo, no se puede dividir por algunos números. División es igual a multiplicar por el inverso y el inverso de un numero $x$ es un número $y (= x^{-1})$ tal que $xy = 1$. En la notación de aritmética modular, trabajando, por ejemplo, módulo 30:

$$\begin{aligned}
7^{-1} &\equiv 13 \pmod{30} \text{ porque } 7 \cdot 13 = 91 \equiv 1 \pmod{30} \\
1^{-1} &\equiv 1 \pmod{30} \text{ porque } 1 \cdot 1 \equiv 1 \pmod{30} \\
29^{-1} &\equiv 29 \pmod{30} \text{ porque } 29 \equiv -1 \pmod{30} \text{ y } (-1)^2 \equiv 1 \pmod{30}
\end{aligned}$$

Usando fuerza bruta se puede ver que 2 no tiene inverso. En general, cuando $\mathrm{mcd}(x, 30) \neq 1$, $x$ no tiene un inverso módulo 30. Este ejemplo es una instancia de un fenómeno bastante general:

**Proposición.** *Sea $n$ un entero positivo y sea $a$ y $b$ dos enteros. Entonces la ecuación*

$$ax \equiv b \pmod{n}$$

*tiene una solución única si y solo si* $\mathrm{mcd}(a, n) = 1$. *Y esta solución es* $a^{-1}b \pmod{n}$.

Ahora se puede terminar con el análisis de lo que pasa cuando uno escribe un símbolo mal en el ISBN. Si el $x_i$ de un número de ISBN se cambia por un $y_i$ entonces la suma $\sum_{k=1}^{10} k x_k$ cambia por $i(y_i - x_i)$. ¿La suma

$$S' = 1 \cdot x_1 + 2 \cdot x_2 + \cdots + i \cdot y_i + \cdots 9 \cdot x_9$$

puede ser igual al símbolo verificador módulo 11? Para fijar ideas, sea $c$ el símbolo verificador. Entonces esta suma debería ser igual a $c$ módulo 11 y la suma $S = \sum_{k=1}^{9} k x_k$ también debería ser igual a $c$ módulo 11. Entonces la diferencia $S' - S$ debería ser igual a 0 módulo 11. La diferencia $S' - S$ es $i(y_i - x_i)$ y la pregunta es: ¿Esta diferencia puede ser 0 módulo 11?

Supongamos que

$$i(y_i - x_i) \equiv 0 \pmod{11}.$$

Como $1 \leq i < 10$, el inverso $i^{-1}$ de $i$ existe y podemos multiplicar cada lado por $i^{-1}$:

$$i^{-1} \cdot i(y_i - x_i) \equiv i^{-1} \cdot 0 \pmod{11}$$
$$(y_i - x_i) \equiv 0 \pmod{11}.$$

Como $0 \leq y_i, x_i \leq 9$, esta última ecuación implica que $x_i = y_i$. Entonces se puede concluir que si uno cambia exactamente uno de los primeros nueve símbolos se puede detectar el error.

Si por alguna razón el recipiente del mensaje supiera que hay un solo error en el mensaje, el recipiente pudiera encontrar el $i$ que contiene el error y corregir el mensaje.

*Regresando al cuento de amor...* Miguel supone que hay un sólo error en lo que anotó y usa el código que detecta errores como un código que corrige errores. Encuentra que hay varias maneras de como corregir el número de ISBN para que tenga un X como el símbolo verificador. Va a Amazon y mira cada ISBN que genera usando este método y cuando ingresa el ISBN 849838205X, inmediatamente reconoce la tapa del libro como la misma del libro que tenía el amor de su vida.

Va corriendo a la librería esa misma noche y compra el libro.

El día siguiente en el ómnibus.

Ella: ¡Wow! Tenemos el mismo libro!

Él : ¡Qué casualidad! Me llamo Miguel.

Ella: Me llamo Yésica...

## 2. El lenguaje secreto del amor.

En los siguientes meses el amor de Miguel y Yésica florece. Les encanta pensar y hablar de la matemática. Como suele suceder, el primer amor matemático para ambos era la teoría de números.

**Cifrados de César.** Un cifrado de César es, en un sentido, el cifrado más clásico de todos – clásico no solo porque viene del periodo clásico de la historia pero también su puesto en el panteón de los cifrados es parecido al puesto que el David de Rafael tiene en el panteón de la escultura. Mostramos como funciona usando un ejemplo. Supongamos que tenemos el *texto plano*:

¿SABÍA YO LO QUE ES AMOR? OJOS JURAD QUE NO PORQUE NUNCA HABÍA VISTO UNA BELLEZA ASÍ.

El cifrado de César procede así: pensamos la letra A como 0, B como 1, C como 2, etc, hasta Z como 26. Entonces la primera palabra del texto plano de arriba se *preprocesa* como

19 0 1 9 0

Entonces, para generar el *texto cifrado* se elige una *clave* que, en este caso es un resto módulo 27; es decir un número de 0 a 26. Por ejemplo si se elige la clave 10 el cifrado de César dice que se suma 10 a cada número en el texto plano preprocesado y se anota el resto módulo 27. Para nuestro ejemplo tenemos

19+10 0+10 1+10 9+10 0+10

que son, módulo 27,

2 10 11 19 10.

Con estos números se determina el texto cifrado usando la misma correspondencia de antes:

C J K S J.

Si Miguel manda un mensaje cifrado, Yésica, para poder leer el mensaje en la manera más eficiente posible, tendría que saber la clave. ¿Qué descifrado debería usar? Empieza con el texto cifrado, lo convierta en números, resta la clave módulo 27, y convierta los números que quedan otra vez en letras. Para hacerlo en una manera eficiente, Yésica tiene que saber la clave.

¿Cuáles son algunos de los problemas con este cifrado (¡Hay muchos!)? Para empezar solamente hay 27 claves distintas y, entonces, la clave para el descifrado es fácil de encontrar usando fuerza bruta. Una segunda limitación de este cifrado es que uno puede usar el método llamado 'análisis frecuencial' para determinar la clave. Si uno ve un mensaje cifrado lo primero que haría es contar que letras aparecen y con que frecuencia lo hacen. La letra en el texto cifrado con la frecuencia más alta probablemente corresponde o a la A o a la E y de ahí se calcula la clave fácilmente. Por ejemplo si la Ñ es la letra más común en el texto cifrado, entonces se adivinaría que la clave $\mathsf{c}$ fuera determinado por una de estas relaciones:

$$0 + \mathsf{c} \equiv 14 \pmod{27} \text{ o } 4 + \mathsf{c} \equiv 14 \pmod{27},$$

(el 0 corresponde a la A, el 4 a la E y el 14 a la Ñ). Si el texto plano que resulta no tiene sentido, se adivina una clave nueva y se repita el proceso.

Este cifrado es un cifrado de tipo sustitución *monoalfabética* porque con aplicar la operación de sustitución se conserva siempre a lo largo de todo el mensaje; en oposición a la sustitución polialfabética. Es decir, por ejemplo, cada E en el texto plano se convierte en la misma letra en el texto cifrado.

*Regresando al cuento de amor...* Los jóvenes pasan todo su tiempo juntos, leyendo libros, hablando de la matemática y paseando por la ciudad. Sus padres, pensando que son demasiados jóvenes, comienzan a prohibir que se vean. Como son padres muy controladores, monitorean las interacciones entre nuestros héroes y, entonces, Miguel y Yésica, enfocan todo su poder matemático para desarrollar métodos para comunicarse en secreto. Ya saben un poco del cifrado de César, incluyendo lo fácil que es para romper, y, entonces inventan un cifrado nuevo. Lo llaman el cifrado de Miguésica. Y deciden que, a partir de ese momento van a comunicar solamente en textos usando ese cifrado. Cada día que se ven en el liceo, compartan la clave que van a usar ese día.

**El cifrado Vigenère.** El cifrado que inventaron también se conoce como el cifrado de Vigenère (según la revista *Scientific American*, aún en el año 1917, casi 400 años después de que fue inventado, se pensaba que era imposible descifrar sin saber la clave). El cifrado funciona así. Consideramos el texto plano

LO ÚNICO QUE ME DUELE DE MORIR, ES QUE NO SEA DE AMOR;

y la clave

LIMON

Preprocesamos el texto plano un poco y repetimos la clave debajo del texto plano

LOUNICOQUEMEDUELEDEMORIRESQUENOSEADEAMOR
LIMONLIMONLIMONLIMONLIMONLIMONLIMONLIMON.

Para calcular el texto de cifrado, de acá se procede como en el cifrado de César. Si la $i$-ésima letra corresponde al número $x_i \in \{0, 1, \ldots, 26\}$ y la $i$-ésima letra de la clave repetida es $y_i$, la $i$-ésima letra $z_i$ del texto de cifrado es la que corresponde a $x_i + y_i \pmod{27}$. Haciéndolo así, llegamos al siguiente texto de cifrado:

VWGBUNWCGQWQOGQVQOQWWZUZQDCGQBWDQOOQOWWZ.

¿Cómo es el descifrado de este cifrado? Si uno conoce la clave, el descifrado es muy parecido al descifrado del cifrado de César: usando la notación del párrafo anterior, simplemente se calcula $z_i - y_i \pmod{27}$.

Este cifrado es más difícil de romper que el cifrado de César porque es un cifrado con sustitución polialfabética: cada ocurrencia de la letra 'E', por ejemplo, en el texto plano, no tiene que ser sustituida por la misma letra en el texto cifrado porque depende de la letra de la clave que corresponde.

*Regresando a nuestro cuento de amor...* Los padres de nuestros héroes se enloquecen por no saber lo que está ocurriendo entre Miguel y Yésica. Deciden contratar a los conocidos matemáticos Dr Friedrich Kasiski y Dr Edward Friedman para romper el cifrado.

**Los métodos de Kasiski y de Friedman.** La prueba de Kasiski toma ventaja de que palabras (o partes de palabras) repetidas pueden ser, por casualidad, cifradas por las mismas letras de la clave. Esto implica que habrán repeticiones en el texto cifrado. Entonces, si uno podría determinar lo largo de la clave, el cifrado se convierte en un cifrado de César con período. El método de Kasiski nos da una manera de encontrar cuantas letras tiene la clave.

Por ejemplo, si consideramos el texto cifrado

RFVSM SFIUO MCCMA CZWKO FGYGR JIVBL RVZCU BLDMF
DGEMG WDWWZ VDSJG MFJGT ZZSFM QHZZS GUDSE ODSCV
EIGWP IZQRG AANQR JWVNX QRCOK QYMUL SDYUY YASJL
ZMWTW FNXAT FFKHV DMCZG SIEJF FAPIP IDSFG YOEJW
WKWXE DCWAO ZEIFW OUFSU SSZID WZZGS OQWVG XRUYE
IWSGY MQRFA NWAQF HWNGA EYCJN WAQFS KZCPI JHAAI
MQRFL RBMWK OENMM PCOVR YEXRJ AQUEM YOQNF SSDOK
NFXEU SWYFM JVFFN HPSUW BBFML VFEBM MIEHG AWQWT
CFHHM ZFNKR GQNRB LRUGR RAMFC OEPCL RUYSD OKNOZ
ULSLH GQEDO KLICY VRWFW UIERG UUEXR IFZID XRZKV
YZHFI FRMBM IWLHJ GVFBG FIKYE OEHDQ VTCEB FMWHI
WRRUW KSFRH XEKWW ELMWF MMAUY YASJQ CSRRR WGCCY
VSJRM EYGSJ VIDEC CKQYY EJVGZ VDIJM GICHS VBWYZ
ARUCV RYEXR GSTOM WZBUB LBSIS SPIYS VZDNM RYXOR
LNDEJ DSEYZ XVVSO FAGFB KHMDY DCJRM KSERM YIOSE
GMFJX MVUMR MKSEC UNMFM XCSYK GIFGS GODFR FDNZG
IEHWQ IZHVA GEIMR KSKYY BVVAA BWARD WSZID GFAGN
OZQFF LNFEY GSJVI DECOK FOBII GLVWU SESKQ YXZLZ
YBWAQ FOMAU YEEHW PUBEQ RWPIY TISFQ YDQZQ SECZS
VLLEU ZSPAA FNQVZ CKBGU IEHJN MQPCO ZNVXE SOSFC

```
QPACN RHMFJ CJGIQ RCOUB HFIDD DNWUS ERWFO REEHS
FNUGR VWEGA WLFSN NDEZR GPIYS GCJHH MJLSJ MUPIJ
QGAIO MUOKR UBVFL AZUNE DOKLG MWRZT BLPIU SDNLA
GRZSZ OVIIR WYIES ACKIY DHVGH EIEMX IABUE MMSKI
YEICZ AZJUH FTGAX AHVSK RFMKF JWFYE EJDDN HFEJR
WYUDK RGQIY DHVGZ BDMWH IWFYM KZHSA YZWLT GAXAI
CZSFH AWUOJ NHGRC SUUIP IVGER LMPUO KLWAV RZWFS
KSPCL RXMVV IFNZQ PZQAQ UPWZB FBGNV VSKNZ QPZQA
QUPUL SZNME SEOVB YZXLG ZBLMW USVRF UVZCQ DOQRF
DMRXQ SWFWP YDXVB SQCQZ VBDNH UISZS QYXPR UGSFA
XRGGO LQRLS KGLMW WFWAN QWTCE BOZTR PWYFA RUSDV
HAPRG GAXMW ECKYF MQRBU BHEYJ JGPYE MEQGZ JDIEG
AOFQW VZNVY ZXFSE CCQDR SFGLQ PFGSY UYSJG MFBUQ
ECKQY MQFFN RHHIE
```

y buscamos conjuntos de letras repetidos como los que ilustramos en el texto cifrado. En este texto se pueden encontrar más de 240 conjuntos de tres, cuatro o cinco letras repetidas. Un tabla de ejemplos:

| Conjunto | Separación | Divisores hasta 20 |
|---|---|---|
| HHM | 2, 3, 6, 7, 13, 14 | |
| IYS | 315 | 3, 5, 7, 9, 15 |
| IYDH | 70 | 2, 5, 7 |
| NZQPZ | 21 | 3, 7 |

Suponiendo que los conjuntos repetidos representan la igualdad del texto plano, la primera fila de la tabla indica que la clave es 2, 3, 6, 7, 13, o 14 de largo, la segunda fila que la clave es 3, 5, 7, 9 o 15 de largo. Entonces adivinamos que es 7 de largo porque en todas la filas en tabla y en más de 200 de las 240 filas de la tabla completa, aparece el divisor 7.

Sabiendo el largo de la clave, ahora se rompe el cifrado de César pensando que el texto cifrado está compuesto por una de cada 7 letras.

Para textos planos más cortos se puede usar el método de Friedman basado en el índice de coincidencia. Este índice da una medida de la variación en la distribución de las letras del texto cifrado. Supongamos que el texto plano tiene $N$ letras: hay $\binom{N}{2}$ maneras de elegir dos letras del texto. Sean $f_1, f_2, \ldots, f_{27}$ las frecuencias con que las letras A, B,..., Z aparecen en el texto plano y $p_1, p_2, \ldots, p_{27}$ las probabilidades de que las letras A, B,..., Z aparezcan en el texto. El índice de coincidencia de un texto $\mathcal{X} = x_1x_2x_3\cdots x_N$ es

$$I_c(\mathcal{X}) = \frac{\sum_{i=1}^{27}\binom{f_i}{2}}{\binom{N}{2}} = \frac{\sum_{i=1}^{27} f_i(f_i-1)}{N(N-1)}$$

que se aproxima con

$$\frac{\sum_{i=1}^{27} f_i^2}{N^2}.$$

Es la probabilidad que dos letras elegidas de un texto son iguales.

Si el texto plano es "natural" en el sentido de que la frecuencia $f_1$ de los A en $\mathcal{X}$ es igual a la frecuencia universal de los A en el castellano, esta suma se estima

como

$$\frac{\sum_{i=1}^{27} f_i^2}{N^2} = \sum_{i=1}^{27} \left(\frac{f_i}{N}\right)^2 \approx \sum_{i=1}^{27} p_i^2.$$

Para el castellano, esta invariante "natural" sería $0,075$. El otro extremo de la invariante es cuando las letras son uniformemente elegidas al azar del alfabeto de 27 letras. En ese caso el índice sería $1/27 \approx 0,037$. La observación clave en lo que sigue es si el cifrado era monoalfabético, se conservaría el índice de coincidencia: el índice del texto del texto plano es igual al índice del texto cifrado. En particular, el texto cifrado bajo un cifrado de César debería tener un índice de coincidencia muy cerca a $0,075$.

Ahora, para adivinar el largo de la clave se hace lo siguiente. Se repartan las letras $x_1, \ldots, x_N$ del texto cifrado entre $r$ columnas ($r$ será el largo de la clave). Entonces hay más o menos

$$r\binom{N/r}{2} = \frac{N(N/r-1)}{2}$$

maneras de elegir dos letras de la misma columna y hay

$$\frac{r \cdot (N/r) \cdot (N - N/r)}{2} = \frac{N(N - N/r)}{2}$$

maneras de elegir dos letras de distintas columnas.

Entonces, uno esperaría tener

$$E = (0,075) \cdot \left(\frac{N(N/r-1)}{2}\right) + (0,037) \cdot \left(\frac{N(N-N/r)}{2}\right)$$

pares iguales y la probabilidad de elegir un par de letras iguales sería

$$\frac{E}{\binom{N}{2}} = \frac{1}{r(N-1)}(0,038N + r(0,037N - 0,075)).$$

Pero el resultado de esta cuenta también es el índice de coincidencia del texto. Entonces,

$$I_c(\mathcal{X}) \approx \frac{1}{r(N-1)}(0,038N + r(0,037N - 0,075))$$

y de aquí se puede resolver para la $r$.

*Regresando a nuestro cuento de amor...* Con la ayuda de los terribles Dres Kasiski y Friedman, los padres ya pueden romper toda la comunicación entre Miguel y Yésica. Tienen una intervención familiar y los padres de Miguel anuncian que están mandando Miguel al exterior para terminar el liceo. Miguel y Yésica piden media hora más, para tener una última oportunidad para estar juntos. Los padres dicen que sí.

Miguel y Yésica van al cuarto de Miguel y empiezan a planear como van a seguir en contacto sin que su padres se enteren. Se dan cuenta de dos cosas: necesitan un cifrado nuevo y una nueva manera de como compartir claves ya que nunca se van a ver y ya que sus padres van a monitorear todos los mensajes no cifrados.

Yésica dice que ella tiene un método para compartir claves.

**Diffie-Hellman.** Este protocolo criptográfico permite que dos personas desconocidas una para la otra puedan establecer una clave secreta y compartida. Mostramos como funciona con un ejemplo, pero primero tenemos que hablar un poco de multiplicación módulo un primo $p$.

Consideramos el conjunto $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ de enteros módulo $p$. Todo elemento de este conjunto menos cero tiene un recíproco multiplicativo (porque $p$ es primo y todo $n$ positivo menor que $p$ es coprimo con $p$). El conjunto $\mathbb{Z}_p^\times = \{1, \ldots, p-1\}$ tiene la estructura de un *grupo*: hay una operación binaria (en nuestro caso multiplicación módulo $p$ denotada $\cdot$) tal que:

1. existe un elemento nulo $e$ tal que $a \cdot e \equiv e \cdot a \equiv a \pmod{p}$ para todo $a \in \mathbb{Z}_p^\times$ (en nuestro caso $e = 1$);
2. la operación binaria es asociativa (o sea $a \cdot (b \cdot c) \equiv (a \cdot b) \cdot c \pmod{p}$ para toda $a, b, c \in \mathbb{Z}_p^\times$;
3. cada elemento de $\mathbb{Z}_p^\times$ tiene un recíproco multiplicativo (o sea, para cada $a \in \mathbb{Z}_p^\times$ existe una $b \in Z_p^\times$ tal que $a \cdot b \equiv 1 \pmod{p}$.

Se dice que el elemento $g \in \mathbb{Z}_p^\times$ *genera* el grupo si cada elemento del grupo se puede escribir como una potencia de $g$. En este caso particular tal $g$ también se llama una *raíz primitiva módulo* $p$.

Miguel y Yésica eligen y publican dos números: un primo $p$ y una raíz primitiva $g$ módulo $p$. Miguel elige un entero $a$ al azar y calcula $u \equiv g^a \pmod{p}$; manda $u$ a Yésica. Yésica elige un entero $b$ al azar y calcula $v \equiv g^b \pmod{p}$; manda $v$ a Miguel. Yésica, entonces, puede calcular la clave $k \equiv u^b \equiv (g^a)^b \pmod{p}$ y Miguel también: $k \equiv v^a \equiv (g^b)^a \pmod{p}$. Ahora tienen la misma clave.

Si sus padres quisieran conocer la clava precisan $a$ o $b$. Sin esos datos, tendrían que resolver el *problema del logaritmo discreto*, un problema para cual no se conoce un algoritmo que lo resuelva en una cantidad de tiempo razonable. El problema es: encontrar $x$ dado $y, g, p$ y $y \equiv g^x \pmod{p}$. El algoritmo de Shanks para resolver este problema usando un primo con 300 dígitos y $a$ y $b$ con 100 dígitos tomaría más tiempo que la edad del universo para romper el protocolo.

*Regresando a nuestro cuento de amor...* Ahora que tienen el método para intercambiar claves, empiezan a discutir el cifrado. Pero en ese momento los padres de Yésica entran y se la llevan. Y Miguel se queda ahí, sentado, sin manera de comunicarse con el amor de su vida.

Nuestros héroes siguen comunicándose en claro, sin un cifrado, pero no pueden hablar como quieren. Hay cosas que solo quieren compartir con el otro, pero ya no pueden. Los padres de Yésica siguen leyendo su mail. Entonces, con la comunicación sofocada, con tiempo, la intimidad de su amor decae. Siguen adelante con sus vidas separadas, trabajando en la matemática y siempre pensando en el otro.

## 3. Como comparar secretamente los amores...

Después de la abrupta separación de nuestros héroes, se dedican independientemente a la matemática en general y la criptografía en particular. Algunas de las cosas que aprendieron, las ilustramos acá.

**Grupos.** Antes de introducir al nuevo sistema criptográfico, se precisan unas ideas básicas. Pensamos de vuelta en equivalencia módulo $n$ y un poco más acerca de grupos. En particular, se sabe que

- un entero $m$ es equivalente módulo $n$ a uno de los enteros $\{0, 1, \ldots, n-1\}$;
- la aritmética funciona en una manera razonable: si $a \equiv b \pmod{n}$ y $c \equiv d \pmod{n}$, sabemos
  - $a + c \equiv b + d \pmod{n}$;
  - $a - c \equiv b - d \pmod{n}$; y
  - $a \times c \equiv b \times d \pmod{n}$;
- la ecuación $ax \equiv b \pmod{n}$ tiene una solución si y solo si $\mcd(a, n) = 1$.

Un grupo es un par de objetos $(G, \star)$ donde $G$ es un conjunto y $\star$ es una operación binaria que cumple

1. [**Clausura**] para toda $x, y \in G$, $x \star y \in G$;
2. [**Elemento nulo**] existe un elemento $e \in G$ tal que $e \star x = x \star e = x$ para todo $x \in G$;
3. [**Asociatividad**] para toda $x, y, z \in G$, $x \star (y \star z) = (x \star y) \star z$;
4. [**Inversos**] para toda $x \in G$, existe $y \in G$ tal que $x \star y = e = y \star x$.

*Dos ejemplos relevantes.* Sean $\mathbb{Z}/n\mathbb{Z}$ el conjunto $\{0, 1, 2, \ldots, n-1\}$ y $\star$ la operación de sumar módulo $n$. Entonces: sabemos de los hechos de arriba que la operación es cerrada; sabemos que $0 \in \mathbb{Z}/n\mathbb{Z}$ sirve como el elemento nulo; una cuenta fea nos deja concluir que la operación es asociativa; y para cada $x$ el elemento $-x \pmod{n}$ es el inverso de $x$.

Sean $(\mathbb{Z}/n\mathbb{Z})^\times = \{x \in \mathbb{Z}/n\mathbb{Z} : \mcd(x, n) = 1\}$ y $\star$ la operación de multiplicar módulo $n$. Entonces: la operación es cerrada porque el producto de dos números coprimos con $n$ es coprimo con $n$; el elemento $1 \in (\mathbb{Z}/n\mathbb{Z})^\times$ es el elemento nulo; una cuenta fea nos dea concluir que la operación es asociativa; y para cada $x$ coprimo con $n$ existe un elemento $y$ tal que $xy \equiv 1 \pmod{n}$.

*La función $\phi$ de Euler.* Varias veces vamos a tener que hablar del tamaño del conjunto $(\mathbb{Z}/n\mathbb{Z})^\times$. Definimos la función

$$\phi : \mathbb{Z}^+ \to \mathbb{Z}^+$$

por $\phi(n) = \#\{x \in \mathbb{Z}/n\mathbb{Z} : \mcd(x, n) = 1\}$. Se puede probar lo siguiente:

**Lema.** *La función $\phi$ es multiplicativa; o sea, si $\mcd(m, n) = 1$, tenemos $\phi(mn) = \phi(m)\phi(n)$. También, si $n$ se factoriza como $n = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}$, tenemos*

$$\phi(n) = n \prod_{p \mid n} \left(1 - \frac{1}{p}\right).$$

Obviamente, $\phi(n)$ es el tamaño del conjunto $(\mathbb{Z}/n\mathbb{Z})^\times$.

*El algoritmo extendido de Euclides.* Para el grupo $(\mathbb{Z}/n\mathbb{Z}, +)$ encontrar el elemento inverso de un elemento dado es fácil y obvio de hacer; para el grupo $((\mathbb{Z}/n\mathbb{Z})^\times, \times)$ también es fácil de hacer pero no tan obvio. Arrancamos con la identidad de Bezout: dado dos enteros $x, y$, existen dos enteros $a, b$ tal que

$$ax + by = \mcd(x, y).$$

Supongamos que queremos encontrar el inverso (multiplicativo) de $x$ módulo $n$; o sea, dado $x$ y $n$ coprimos, ¿cómo encontramos la $y$ tal que $xy \equiv 1 \pmod{n}$?; o sea, ¿como encontramos el entero $k$ tal que $xy + kn = 1 (= \mcd(x, n))$? Según la identidad de Bezout tal $k$ e $y$ existen pero ahora hablamos de como encontrarlo.

**Algoritmo 1.** *Supongamos que* $x$ *e* $y$ *son enteros y sea* $m = \mathrm{mcd}(x, y)$. *Este algoritmo encuentra enteros* $a$ *y* $b$ *tal que* $xa + yb = m$.

1. ***[Inicializar]*** *Sean* $a = 1$, $b = 0$, $r = 0$, $s = 1$.
2. ***[¿Terminado?]*** *Si* $y = 0$, *sea* $m = x$ *y terminar.*
3. ***[Cociente y resto]*** *Escribir* $x = qy + c$ *con* $0 \leq c < y$.
4. ***[Shiftear]*** *Sea* $(x, y, r, s, a, b) = (y, c, a - qr, b - qs, r, s)$ *y seguir a paso (2).*

Ahora un ejemplo:

**Ejemplo 3.1.** Resolvemos la ecuación $17x \equiv 1 \pmod{17}$. Primero usamos el algoritmo recién descrito para encontrar $a$ y $b$ tal que $17x + 61y = 1$:

$$\begin{aligned} \underline{61} &= 3 \cdot \underline{17} + \underline{10} \\ \underline{17} &= 1 \cdot \underline{10} + \underline{7} \\ \underline{10} &= 1 \cdot \underline{7} + \underline{3} \\ \underline{7} &= 2 \cdot \underline{3} + \underline{1}. \end{aligned}$$

Trabajando con estas cuentas pero trabajando al revés, tenemos,

$$\begin{aligned} \underline{1} &= \underline{7} - 2 \cdot \underline{3} \\ &= \underline{7} - 2 \cdot (\underline{10} - 1 \cdot \underline{7}) \\ &= 3 \cdot \underline{7} - 2 \cdot \underline{10} \\ &= 3 \cdot (\underline{17} - 1 \cdot \underline{10}) - 2 \cdot (\underline{61} - 3 \cdot \underline{17}) \\ &= 9 \cdot \underline{17} - 2 \cdot \underline{61} - 3 \cdot \underline{10} \\ &= 9 \cdot \underline{17} - 2 \cdot \underline{61} - 3 \cdot (\underline{61} - 3 \cdot \underline{17}) \\ &= 18 \cdot \underline{17} - 5 \cdot \underline{61}. \end{aligned}$$

O sea

$$1 \equiv 18 \cdot 17 \pmod{61}.$$

En particular, esto quiere decir que el inverso multiplicativo de 17 módulo 61 es 18 módulo 61.

Calcular el inverso así es fácil (o, mejor dicho, rápido) por este teorema de Lamé:

**Teorema.** *La cantidad de pasos requerido por el algoritmo de arriba es menor o igual a 5 por el número de dígitos en* $\max\{x, y\}$.

*El pequeño teorema de Fermat.* Este teorema tiene muchas aplicaciones por todos lados de la matemática.

**Teorema.** *Sea* $p$ *un primo y* $a$ *un entero tal que* $\mathrm{mcd}(a, p) = 1$. *Entonces* $a^{p-1} \equiv 1 \pmod{p}$.

*Demostración.* Miramos la sucesión $a, 2a, 3a, \ldots, (p-1)a$. Ninguno de los enteros son congruentes módulo $p$: si lo fueran, $ma \equiv na \pmod{p}$ y, como $a$ es coprimo con $p$ se puede "dividir" por $a$ y deducir que $m \equiv n \pmod{p}$, un absurdo. Además, como $a$ y $p$ son coprimos, ninguno de los números $ma$ es equivalente a cero módulo $p$. Entonces

$$\begin{aligned} 1 \cdot 2 \cdot \cdots \cdot (p-1) &\equiv a \cdot 2a \cdot \cdots \cdot (p-1)a \pmod{p} \\ (p-1)! &\equiv a^{p-1}(p-1)! \pmod{p}. \end{aligned}$$

Como $p$ y $(p-1)!$ son coprimos, se puede dividir por $(p-1)!$ y llegamos a

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

□

**El sistema RSA.** El sistema criptográfico asimétrico RSA fue desarrollado en 1977 por Rivest, Shamir y Adleman. Se llama asimétrico porque de los dos participantes, solamente uno tiene control de las claves. Por ejemplo si el sistema es de Yésica, ella genera dos claves: una clave pública para que cualquiera otra persona puede comunicarse con Yésica en una forma privada y, para cada clave pública, genera una clave privada que ella usa para el descifrado.

*Los detalles del sistema RSA.* La idea fundamental que forma la base de RSA es la construcción por Yésica de una función invertible

$$E : X \to X$$

tal que cualquiera persona puede calcular $E(x)$ pero solamente Yésica puede calcular $E^{-1}(x)$. Yésica construye tal función de esta manera:

1. Yésica elige dos primos grandes $p$ y $q$; calcula $n = pq$.
2. Yésica puede calcular
$$\phi(n) = (p-1)(q-1).$$
3. Yésica elige un entero $e$ al azar tal que
$$1 < e < \phi(n) \text{ y } \operatorname{mcd}(e, \phi(n)) = 1.$$
4. Yésica calcula una solución $x = d$ a la congruencia
$$ex \equiv 1 \pmod{\phi(n)}.$$
5. Finalmente, Yésica define una función $E : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ por
$$E(x) = x^e \in \mathbb{Z}/n\mathbb{Z}.$$

Para mandar un mensaje a Yésica se procede así. Se traduce el mensaje, de alguna manera, a una sucesión de números módulo $n$:

$$m_1, \ldots, m_r \in \mathbb{Z}/n\mathbb{Z}$$

y se manda

$$E(m_1), \ldots, E(m_r)$$

o sea,

$$m_1^e \pmod{n}, \ldots, m_r^e \pmod{n}.$$

Cuando los $E(m_i)$ llegan a Yésica, ella puede calcular $E^{-1}(m) = m^d \pmod{n}$. El hecho que el inverso se calcula usando el $d$ viene del siguiente resultado.

**Proposición.** *Sea $n$ un entero que es un producto de primos distintos y sean $d, e \in \mathbb{Z}_+$ tal que $p - 1 \mid de - 1$ para cada primo que divide a $n$. Entonces $a^{de} \equiv a \pmod{n}$ para todo $a \in \mathbb{Z}$.*

*Demostración.* Como $n | a^{de} - a$ si y solo si para cada $p \mid n$ tenemos que $p \mid a^{de} - a$, es suficiente probar que $a^{de} \equiv a \pmod{p}$ para cada $p \mid n$. Entonces hay dos casos: (1) $\operatorname{mcd}(a, p) \neq 1$: en este caso $a \equiv 0 \pmod{p}$ y, entonces, $a^{de} \equiv a \pmod{p}$; (2) $\operatorname{mcd}(a, p) = 1$: en este caso, el pequeño teorema de Fermat afirma que $a^{p-1} \equiv 1 \pmod{p}$. Ahora, como para algún $k$ tenemos $k(p-1) = de - 1$, se ve que

$$1 \equiv 1^k \equiv (a^{p-1})^k \equiv a^{de-1} \pmod{p}$$

Multiplicando cada lado por $a$ nos da

$$a^{de} \equiv a \pmod{n}.$$

□

Entonces, para calcular el descifrado se calcula

$$E^{-1}(m) \equiv E(m_i)^d \equiv (m_i^e)^d \equiv m_i \pmod{n}.$$

*Convirtiendo un texto plano a un número.* Hay varias maneras de hacerlo y para hacerlo de verdad es mucho más complicado de los que vamos a ver ahora. Por ejemplo, hay que "rellenar" el texto plano con caracteres aleatorios para evitar ciertos ataques.

Sea $s$ una cadena de letras mayúsculas y espacios tal que no empieza con un espacio. Se convierte $s$ a un número de esta manera: un espacio corresponde a 0, la letra A a 1, ..., Z a 27, y escribimos el número en base 28. Entonces

YESICA

corresponde a

$$28^5 \cdot 26 + 28^4 \cdot 5 + 28^3 \cdot 20 + 28^2 \cdot 9 + 28^1 \cdot 3 + 28^0 \cdot 1 = 450989029.$$

Para ir al revés, se tiene que dividir por potencias sucesivas de 28.

Si $28^k \leq n$ se puede cifrar cualquier sucesión de $k$ letras tal que el resultado, después de aplicar el método de arriba, es menor o igual a $n$. Entonces, si se puede cifrar números de tamaño no mayor que $n$, se tiene que separar el mensaje original en bloques $m_i$ de tamaño no mayor que $\log_{28}(n)$.

*Un ejemplo completo de RSA.*

1. Sean

$$p = 5032942093845743985781 \text{ y } q = 14032942093845743985769.$$

2. Entonces

$$\begin{aligned}\phi(n) &= \phi(p)\phi(q)\\ &= (p-1)(q-1)\\ &= 70626984964616077533542398459436891914379040.\end{aligned}$$

3. Elegir al azar un $e < \phi(n)$:

$$e = 69418451666598544362041409492071945586962923$$

y publicar la clave pública: $(e, \phi(n))$.

4. Calcular el $d \pmod{\phi(n)}$ tal que $e \cdot d \equiv 1 \pmod{\phi(n)}$:

$$d = 23617223401654479430458819886672049101674307.$$

5. Cifrar la letra X, que corresponde al número 25:

$$E(X) \equiv 25^e \equiv 16828894343284430278543573571004692857545385 \pmod{\phi(n)}$$

6. Para descifrar, se calcula $E(X)^d$ y se puede ver (usando una calculadora) que

$$E(X)^d = X.$$

*Exponenciación rápida.* Supongamos que queremos calcular $2^{17}$. El algoritmo obvio de como hacerlo es multiplicar dos por dos 16 veces:

$$2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2$$

Mejor sería

$$\left(\left((2 \cdot 2)^2\right)^2\right)^2 \cdot 2$$

que requiere cinco multiplicaciones solamente. Además si uno estuviera calculando el resto de la potencia módulo $n$, se reduciría módulo $n$ después de cada producto. Entonces se calcula $b^m \pmod n$ de esta manera:

**Algoritmo 2.** *Dado enteros $b$, $m$ y $n$, este algoritmo calcula $x = b^m \pmod n$ en una manera eficiente.*

1. ***[Inicializar]*** *Poner $x = 1$ y escribir $m$ en base 2: $(m)_2$.*
2. ***[Seguir?]*** *$m > 0$? Si no, devolver el valor de $x$.*
3. ***[Caso impar]*** *Si el último dígito de $(m)_2$ es 1, poner $x = x \cdot b \pmod n$.*
4. ***[Move a la izquierda]*** *Sacar el último dígito de $(m)_2$ y hacer el cambio correspondiente a $m$ (dividir por 2, o restar 1 y después dividir por 2).*
5. ***[Cuadrar la base]*** *Poner $b = b^2 \pmod n$.*
6. ***[Continuar]*** *Regresar a paso (2).*

**Hay que tener cuidado con las claves.** Describimos dos ejemplos donde uno puede, sin tanto esfuerzo, encontrar la clave privada.

*Factorizando $n$ dado $\phi(n)$.* Sabemos que $\phi(n) = pq - p - q + 1$ y que $pq = n$. Entonces, si consideramos el polinomio

$$x^2 - (p + q)x + pq = x^2 - (n - \phi(n) + 1)x + n$$

se ve que las raíces del polinomio son $p$ y $q$ y, usando propiedades de la función $\phi$, sabemos los coeficientes del polinomio. Las raíces se pueden encontrar usando la fórmula cuadrática.

*Factorizando $n$ dado que $p$ y $q$ están cercas.* Éste es el método de factorización de Fermat: sean $p > q$ y $n = pq$. Entonces

$$n = \left(\frac{p+q}{2}\right)^2 - \left(\frac{p-q}{2}\right)^2.$$

Ahora si $p \approx q$, la cantidad $\left(\frac{p-q}{2}\right)^2$ es muy pequeña, la cantidad $\left(\frac{p+q}{2}\right)$ es aproximadamente el tamaño de $\sqrt{n}$ y $t^2 - n = s^2$ es un cuadrado perfecto. Entonces probamos

$$t = \lceil\sqrt{n}\rceil,\ t = \lceil\sqrt{n}\rceil + 1,\ \lceil\sqrt{n}\rceil + 2, \ldots$$

hasta que $t^2 - n$ sea un cuadrado. Entonces

$$p = t + s \text{ y } q = t - s.$$

*Regresando a nuestro cuento de amor...* Después de haber pasado tantos años separados, después que ambos habían ganado sus doctorados en matemática, se ven en una fiesta tirada por un amigo común. Cuando se ven, es como si el tiempo no se había movido en los años desde la última vez que pasaron unos momentos juntos. Miguel siente que sigue enamorado de Yésica pero no quiere anunciar que sigue con estos sentimientos sin saber que Yésica siente lo mismo. Entonces pasa un rato pensando en como resolver el problema de saber si Yésica está enamorado de él en una manera que se quedaría sin pedir la ayuda de ninguna otra persona.

**Comunicación segura entres varios partidos.** Supongamos que hay dos millonarios y quieren saber cual de los dos es el más rico pero sin revelar la cantidad de plata que tienen ni al otro ni a cualquier otra persona. ¿Cómo se puede resolver la duda sin revelar nada?

Millonario 1 tiene I millones de dólares y Millonario 2 tiene J millones de dólares. Quieren averiguar si $I \leq J$ pero al final del protocolo ninguno de los dos debería haber aprendido nada con respeto a la riqueza de la otra persona salvo la información implicada por $I \geq J$. Para simplicidad, supongamos que $I, J \in \{1, \ldots, 10\}$.

Supongamos que el millonario 2 tiene una clave privada y pública de RSA: $(d, n)$ y $(e, n)$, respectivamente y usa la función $E_2$ para cifrar y $D_2$ para descifrar. El protocolo es el siguiente:

1. Millonario 1 elige un número grande $x$ y lo cifra usando el cifrado de Millonario 2: $c = E_2(x)$.
2. Millonario 1 calcula $c - I$ y manda el resultado a millonario 2.
3. Millonario 2 calcula $y_u = D_2(c - I + u)$ para cada $1 \leq u \leq 10$.
4. Millonario 2 elige un primo $p$ de tamaño aproximadamente $x/2$ tal que $|z_u - z_v| \geq 2$ para $u \neq v$ $(u, v \in \{1, \ldots, 10\})$ donde $z_u \equiv y_u \pmod{p}$ para $u \in \{1, \ldots, 10\}$.
5. Millonario 2 manda a millonario 1 esta sucesión de números en este orden:
$$z_1, z_2, \ldots, z_J, z_{J+1} + 1, z_{J+2} + 1, \ldots, z_{10} + 1, p.$$
6. Millonario 1 verifica si el I-ésimo número de esta sucesión es equivalente a $x$ módulo $p$. Si lo es, deduce que $I \leq J$; sino, deduce que $I > J$.

*Regresando a nuestro cuento de amor...* Miguel se acuerda del protocolo para resolver el problema de los dos millonarios, se acerca a Yésica, y le dice:

Él: Elegí un número: 3 si sigues enamorado de mí y 5 si no. Yo elijo 4 si estoy enamorado de ti y 2 si no. Sin contarme tu número vamos a deducir que hacer con lo que hay entre nosotros dos...

Ella: ¿Cómo vas a poder hacer eso?

Él: Si el número que estás pensando es mayor de la que yo estoy pensando, seguimos enamorados. ¿Lo probamos?

Instituto de Matemática y Estadística Rafael Laguardia
Universidad de la República, Uruguay
*E-mail address*: nryan@fing.edu.uy