Tesis de Licenciatura en Ciencias de la Computación

# Números Muy Normales

Lucas Puterman

Directora: Verónica Becher
Departamento de Computación
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
20 de Noviembre, 2019

# Sobre secuencias aleatorias


Supongamos que tiramos una moneda infinitas veces y anotamos un $1$ cada vez que sale cara y $0$ cada vez que sale ceca ¿Cuáles de estas secuencias es creíble que sea el resultado de este experimento?

# Sobre secuencias aleatorias


Supongamos que tiramos una moneda infinitas veces y anotamos un $1$ cada vez que sale cara y $0$ cada vez que sale ceca ¿Cuáles de estas secuencias es creíble que sea el resultado de este experimento?

- $11111111111111111111111111111111111\ldots$

# Sobre secuencias aleatorias

Supongamos que tiramos una moneda infinitas veces y anotamos un $1$ cada vez que sale cara y $0$ cada vez que sale ceca ¿Cuáles de estas secuencias es creíble que sea el resultado de este experimento?

- $11111111111111111111111111111111111\ldots$
- $01001000100001000001000000100000001\ldots$

# Sobre secuencias aleatorias


Supongamos que tiramos una moneda infinitas veces y anotamos un $1$ cada vez que sale cara y $0$ cada vez que sale ceca ¿Cuáles de estas secuencias es creíble que sea el resultado de este experimento?

- $11111111111111111111111111111111\ldots$
- $01001000100001000001000000100000001\ldots$
- $01010101010101010101010101010101010\ldots$

# Sobre secuencias aleatorias


Supongamos que tiramos una moneda infinitas veces y anotamos un $1$ cada vez que sale cara y $0$ cada vez que sale ceca ¿Cuáles de estas secuencias es creíble que sea el resultado de este experimento?

- $11111111111111111111111111111111111\dots$
- $01001000100001000001000000100000001\dots$
- $01010101010101010101010101010101010\dots$
- $10000110001010001110010010110011010\dots$

# Secuencias normales

Podemos pensar que en una secuencia aleatoria no hay ningún patrón de $\ell$ símbolos que sea más frecuente que otro.

# Secuencias normales

Podemos pensar que en una secuencia aleatoria no hay ningún patrón de $\ell$ símbolos que sea más frecuente que otro.

## Definición

Notamos $|u|_v$ a la *cantidad de ocurrencias* de la palabra $v$ dentro de la palabra $u$.

# Secuencias normales

Podemos pensar que en una secuencia aleatoria no hay ningún patrón de $\ell$ símbolos que sea más frecuente que otro.

## Definición

Notamos $|u|_v$ a la *cantidad de ocurrencias* de la palabra $v$ dentro de la palabra $u$.

Además, notamos $u[i, j]$ a la subsecuencia de $u$ formada tomando todos los símbolos entre el $i$ y el $j$ inclusive.

# Secuencias normales

Podemos pensar que en una secuencia aleatoria no hay ningún patrón de $\ell$ símbolos que sea más frecuente que otro.

## Definición

Notamos $|u|_v$ a la *cantidad de ocurrencias* de la palabra $v$ dentro de la palabra $u$.

Además, notamos $u[i, j]$ a la subsecuencia de $u$ formada tomando todos los símbolos entre el $i$ y el $j$ inclusive.

Nos gustaría que para un prefijo de una secuencia aleatoria suficientemente grande, la cantidad de ocurrencias de cada palabra de cierta longitud sea casi la misma.

# Secuencias normales

## Definición (Borel, 1909)

Dado un alfabeto $A$ y alguna secuencia infinita $u \in A^\omega$, decimos que $u$ es *simplemente normal para la longitud* $\ell$ si cada secuencia $v$ de longitud $\ell$ verifica que

$$\lim_{n\to\infty} \frac{|u[1,\ell n]|_v}{n} = \frac{1}{|A|^\ell}.$$

# Secuencias normales

### Definición (Borel, 1909)

Dado un alfabeto $A$ y alguna secuencia infinita $u \in A^\omega$, decimos que $u$ es *simplemente normal para la longitud* $\ell$ si cada secuencia $v$ de longitud $\ell$ verifica que

$$\lim_{n \to \infty} \frac{|u[1, \ell n]|_v}{n} = \frac{1}{|A|^\ell}.$$

Decimos que $u$ es *normal* si es simplemente normal para toda longitud $\ell \in \mathbb{N}$.

# Secuencias normales

## Definición (Borel, 1909)

Dado un alfabeto $A$ y alguna secuencia infinita $u \in A^\omega$, decimos que $u$ es *simplemente normal para la longitud* $\ell$ si cada secuencia $v$ de longitud $\ell$ verifica que

$$\lim_{n\to\infty} \frac{|u[1,\ell n]|_v}{n} = \frac{1}{|A|^\ell}.$$

Decimos que $u$ es *normal* si es simplemente normal para toda longitud $\ell \in \mathbb{N}$.

## Problema (Borel, 1909)

*Encontrar ejemplos naturales de secuencias normales.*
*Decidir si la representación en base $b$ de $\pi$, $e$ ó $\sqrt{2}$ es normal.*

# La secuencia de Champernowne

## Problema

*Encontrar algún ejemplo explícito de una secuencia normal.*

# La secuencia de Champernowne

### Problema

*Encontrar algún ejemplo explícito de una secuencia normal.*

### Teorema (Champernowne, 1933)

*La secuencia*

$$1234567891011121314151617181920\ldots$$

*es normal sobre el alfabeto* $A = \{0, 1, \ldots, 9\}$.

# $champ$, La secuencia que usaremos

**Teorema** (Bugeaud, 2012)

*Sea $A$ un alfabeto. Llamamos $X(n)$ a la concatenación de todas las palabras de lomgitud $n$ formadas por símbolos de $A$ en orden lexicográfico.*

*La palabra infinita $X(1)X(2)\ldots$ es normal en el alfabeto $A$*

# $champ$, La secuencia que usaremos

**Teorema** (Bugeaud, 2012)

*Sea $A$ un alfabeto. Llamamos $X(n)$ a la concatenación de todas las palabras de lomgitud $n$ formadas por símbolos de $A$ en orden lexicográfico.*
*La palabra infinita $X(1)X(2)\ldots$ es normal en el alfabeto $A$*

En particular, nosotros vamos a usar el alfabeto $A = \{0, 1\}$ Entonces, por ejemplo:

$$X(2) = 00\,01\,10\,11$$

# $champ$, La secuencia que usaremos

### Teorema (Bugeaud, 2012)

*Sea $A$ un alfabeto. Llamamos $X(n)$ a la concatenación de todas las palabras de lomgitud $n$ formadas por símbolos de $A$ en orden lexicográfico.*
*La palabra infinita $X(1)X(2)\dots$ es normal en el alfabeto $A$*

En particular, nosotros vamos a usar el alfabeto $A = \{0, 1\}$ Entonces, por ejemplo:

$$X(2) = 00\,01\,10\,11$$

Entonces, los primeros símbolos de la secuencia que llamamos $champ$

son:

$$champ = 0\;1\;00\;01\;10\;11\;000\;001\;010\;011\;100\;101\;110\;111\;0000\;0001\;\dots$$

# Supernormalidad

Sea $x$ una secuencia binaria. Sea $A^{\lambda}_{k,n}(x)$ la frecuencia de ocurrencia de las palabras de longitud $n$ que ocurren exactamente $k$ veces comenzando en las primeras $\lfloor \lambda 2^n \rfloor$ posiciones de $x$. Es decir:

$$A^{\lambda}_{k,n}(x) = \frac{\#\{w : |w| = n, |x[1 \ldots \lfloor \lambda 2^n \rfloor]|_w = k\}}{2^n}$$

# Supernormalidad

Sea $x$ una secuencia binaria. Sea $A_{k,n}^{\lambda}(x)$ la frecuencia de ocurrencia de las palabras de longitud $n$ que ocurren exactamente $k$ veces comenzando en las primeras $\lfloor \lambda 2^n \rfloor$ posiciones de $x$. Es decir:

$$A_{k,n}^{\lambda}(x) = \frac{\#\{w : |w| = n, |x[1 \ldots \lfloor \lambda 2^n \rfloor]|_w = k\}}{2^n}$$

## Definición

Sea $\lambda$ un real mayor a cero. Decimos que la secuencia binaria $x$ es $\lambda$-supernormal si para todo entero no negativo $k$ sucede que

$$\lim_{n \to \infty} A_{k,n}^{\lambda}(x) = \frac{e^{-\lambda} \lambda^k}{k!}$$

Decimos que $x$ es supernormal si es $\lambda$-supernormal para todo $\lambda$.

# Un ejemplo para entender mejor

Veamos como ejemplo de juguete si la secuencia finita $x = 10011110$ es supernormal tomando $n = 3$ y $\lambda = 1$.

# Un ejemplo para entender mejor

Veamos como ejemplo de juguete si la secuencia finita $x = 10011110$ es supernormal tomando $n = 3$ y $\lambda = 1$. Las palabras de tamaño 3 que ocurren en $x$ son:

$$100,\ 001,\ 011,\ 111,\ 111,\ 110$$

# Un ejemplo para entender mejor

Veamos como ejemplo de juguete si la secuencia finita $x = 10011110$ es supernormal tomando $n = 3$ y $\lambda = 1$. Las palabras de tamaño 3 que ocurren en $x$ son:

$$100,\ 001,\ 011,\ 111,\ 111,\ 110$$

Si contamos las cantidad de ocurrencias de cada palabra de tamaño 3 tenemos:

| Word | Count |
|---|---|
| 000 | 0 |
| 001 | 1 |
| 010 | 0 |
| 011 | 1 |
| 100 | 1 |
| 101 | 0 |
| 110 | 1 |
| 111 | 2 |

# Un ejemplo para entender mejor

Ahora, veamos las cantidad, las frecuencias y el valor esperado para cada $k$ posible si $x$ fuera 1-supernormal.

# Un ejemplo para entender mejor

Ahora, veamos las cantidad, las frecuencias y el valor esperado para cada $k$ posible si $x$ fuera 1-supernormal.

| $k$ | Count | Frequency | Expected Frequency |
|---|---|---|---|
| 0 | 3 | $\frac{3}{8}$ | $e^{-1}$ |
| 1 | 4 | $\frac{1}{2}$ | $e^{-1}$ |
| 2 | 1 | $\frac{1}{8}$ | $\frac{e^{-1}}{2}$ |
| 3 | 0 | 0 | $\frac{e^{-1}}{3!}$ |
| 4 | 0 | 0 | $\frac{e^{-1}}{4!}$ |
| 5 | 0 | 0 | $\frac{e^{-1}}{5!}$ |
| 6 | 0 | 0 | $\frac{e^{-1}}{6!}$ |
| 7 | 0 | 0 | $\frac{e^{-1}}{7!}$ |
| 8 | 0 | 0 | $\frac{e^{-1}}{8!}$ |

# El resultado de esta tesis

## Teorema

*La noción de supernormalidad es más fuerte que la noción de normalidad. Es decir, los siguientes enunciados son ciertos:*

1. *Sea $x$ una secuencia infinita. Si $x$ es normal, no necesariamente $x$ es supernormal. (normal $\not\Rightarrow$ supernormal )*
2. *Sea $x$ una secuencia infinita. Si $x$ es supernormal, entonces $x$ es normal. (supernormal $\Rightarrow$ normal )*

# ¿Es $champ$ supernormal?

La forma más simple de ver que una secuencia normal no es supernormal es encontrar un ejemplo, y que mejor ejemplo que $champ$.

## Recordemos

*$X(n)$ es la concatenación de todas las palabras de longitud $n$ sobre el alfabeto $A = \{0, 1\}$ en order lexicográfico.*
*Llamamos $champ$ a la concatenación de $X(n)$ para $n = 1, 2, \ldots$*

$$champ = 0\ 1\ 00\ 01\ 10\ 11\ 000\ 001\ 010\ 011\ 100\ 101\ 110\ 111\ 0000\ 0001\ \ldots$$

# ¿Es $champ$ supernormal?

La forma más simple de ver que una secuencia normal no es supernormal es encontrar un ejemplo, y que mejor ejemplo que $champ$.

## Recordemos

*$X(n)$ es la concatenación de todas las palabras de longitud $n$ sobre el alfabeto $A = \{0, 1\}$ en order lexicográfico.*
*Llamamos $champ$ a la concatenación de $X(n)$ para $n = 1, 2, \ldots$*

$$champ = 0\ 1\ 00\ 01\ 10\ 11\ 000\ 001\ 010\ 011\ 100\ 101\ 110\ 111\ 0000\ 0001\ \ldots$$

¿Qué sucede cuando contamos la cantidad de palabras de tamaño $n$ en los primeros $2^n$ símbolos de $champ$?

# ¿Es $champ$ supernormal?

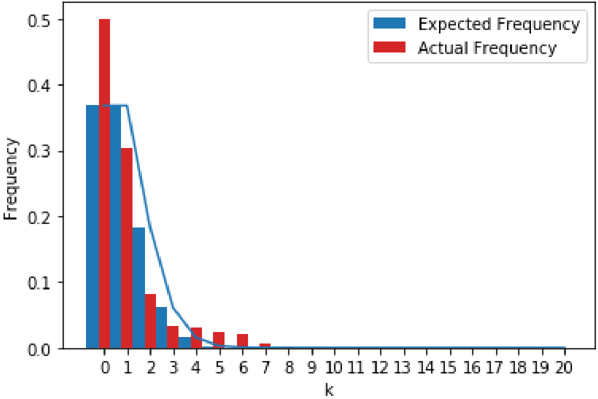


Figura: Frecuencias observadas y esperadas en $champ$ for $n = 16$ y $\lambda = 1$.

# ¿Es *champ* supernormal?

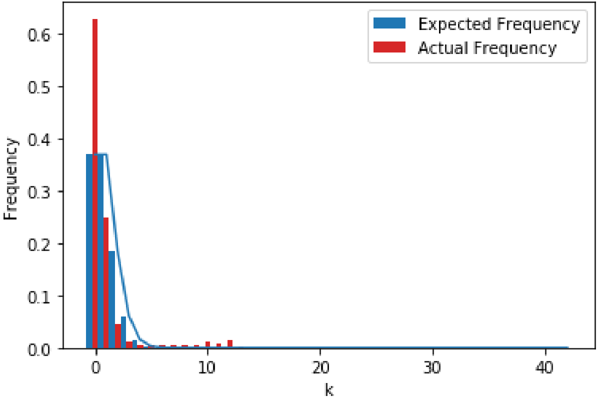


Figura: Frecuencias observadas y esperadas en *champ* for $n = 22$ y $\lambda = 1$.

# Normal $\not\Rightarrow$ Supernormal - Idea de la demostración

- Sabemos que $champ = X(1)X(2)X(3)\dots$

# Normal $\not\Rightarrow$ Supernormal - Idea de la demostración

- Sabemos que $champ = X(1)X(2)X(3)\dots$
- La idea va a ser encontrar un $k$ tal que $X(k)$ sea lo suficientemente grande y este completamente contenido en $champ[1 \dots 2^n]$ y analizar cuales son las palabras distintas que aparecen.

# Normal $\not\Rightarrow$ Supernormal - Idea de la demostración

- Sabemos que $champ = X(1)X(2)X(3)\ldots$
- La idea va a ser encontrar un $k$ tal que $X(k)$ sea lo suficientemente grande y este completamente contenido en $champ[1\ldots 2^n]$ y analizar cuales son las palabras distintas que aparecen.
- Si tomamos $n \geq d + k + 1$ con $k = 2^d$, sucede que $X(k)$ está completamente contenido en $champ[1\ldots 2^n]$.

# Normal $\not\Rightarrow$ Supernormal - Idea de la demostración

- Sabemos que $champ = X(1)X(2)X(3)\ldots$
- La idea va a ser encontrar un $k$ tal que $X(k)$ sea lo suficientemente grande y este completamente contenido en $champ[1\ldots 2^n]$ y analizar cuales son las palabras distintas que aparecen.
- Si tomamos $n \geq d + k + 1$ con $k = 2^d$, sucede que $X(k)$ está completamente contenido en $champ[1\ldots 2^n]$.
- Nos fijamos que pinta tienen las secuencias de longitud $d + k + 1$ que aparecen adentro de $X(k)$ y damos una cota para la cantidad de palabras distintas que pueden aparecer.

# Normal $\not\Rightarrow$ Supernormal - Idea de la demostración

- Sabemos que $champ = X(1)X(2)X(3)\dots$
- La idea va a ser encontrar un $k$ tal que $X(k)$ sea lo suficientemente grande y este completamente contenido en $champ[1 \dots 2^n]$ y analizar cuales son las palabras distintas que aparecen.
- Si tomamos $n \geq d + k + 1$ con $k = 2^d$, sucede que $X(k)$ está completamente contenido en $champ[1 \dots 2^n]$.
- Nos fijamos que pinta tienen las secuencias de longitud $d + k + 1$ que aparecen adentro de $X(k)$ y damos una cota para la cantidad de palabras distintas que pueden aparecer.
- Por último resta ver que la cantidad de palabras que aparecen es mucho menor de lo necesario para que $champ$ sea supernormal.

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$00000000 \qquad 00000001 \qquad 00000010 \qquad 00000011 \qquad 00000100$$

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$(00000000 \qquad 0000)0001 \qquad 00000010 \qquad 00000011 \qquad 00000100$$

- Caso 1:

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$0(0000000 \qquad 00000)001 \qquad 00000010 \qquad 00000011 \qquad 00000100$$

- Caso 1:

$$x = u_1 u_2 \ldots u_k \quad v_1 v_2 \ldots v_d v_{d+1}$$

- Caso 3:

$$x = u_{n+1} u_{n+2} \ldots u_k \quad v_1 v_2 \ldots v_{d+n+1}$$

con $n \in \{1, 2, \ldots, k - d - 2\}$.

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$00(000000 \qquad 000000)01 \qquad 00000010 \qquad 00000011 \qquad 00000100$$

- Caso 1:

$$x = u_1 u_2 \ldots u_k \quad v_1 v_2 \ldots v_d v_{d+1}$$

- Caso 3:

$$x = u_{n+1} u_{n+2} \ldots u_k \quad v_1 v_2 \ldots v_{d+n+1}$$

con $n \in \{1, 2, \ldots, k - d - 2\}$.

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$000(00000 \qquad 0000000)1 \qquad 00000010 \qquad 00000011 \qquad 00000100$$

- Caso 1:

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

- Caso 3:

$$x = u_{n+1} u_{n+2} \dots u_k \quad v_1 v_2 \dots v_{d+n+1}$$

con $n \in \{1, 2, \dots, k - d - 2\}$.

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$0000(0000 \quad 00000001) \quad 00000010 \quad 00000011 \quad 00000100$$

- Caso 1:

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

- Caso 2:

$$x = u_{k-d-1} \dots u_k \quad v_1 v_2 \dots v_k$$

- Caso 3:

$$x = u_{n+1} u_{n+2} \dots u_k \quad v_1 v_2 \dots v_{d+n+1}$$

con $n \in \{1, 2, \dots, k - d - 2\}$.

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$00000(000 \quad 00000001 \quad 0)0000010 \quad 00000011 \quad 00000100$$

- Caso 1:

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

- Caso 2:

$$x = u_{k-d-1} \dots u_k \quad v_1 v_2 \dots v_k$$

- Caso 3:

$$x = u_{n+1} u_{n+2} \dots u_k \quad v_1 v_2 \dots v_{d+n+1}$$

  con $n \in \{1, 2, \dots, k - d - 2\}$.

- Caso 4:

$$x = u_{k-d-1+n} u_{k-d+n} \dots u_k \quad v_1 v_2 \dots v_k \quad w_1 w_2 \dots w_{d+1-n}$$

  con $n \in \{1, 2, \dots, d\}$

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$000000(00 \qquad 00000001 \qquad 00)000010 \qquad 00000011 \qquad 00000100$$

- Caso 1:

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

- Caso 2:

$$x = u_{k-d-1} \dots u_k \quad v_1 v_2 \dots v_k$$

- Caso 3:

$$x = u_{n+1} u_{n+2} \dots u_k \quad v_1 v_2 \dots v_{d+n+1}$$

  con $n \in \{1, 2, \dots, k - d - 2\}$.

- Caso 4:

$$x = u_{k-d-1+n} u_{k-d+n} \dots u_k \quad v_1 v_2 \dots v_k \quad w_1 w_2 \dots w_{d+1-n}$$

  con $n \in \{1, 2, \dots, d\}$

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$0000000(0 \qquad 00000001 \qquad 000)00010 \qquad 00000011 \qquad 00000100$$

- Caso 1:

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

- Caso 2:

$$x = u_{k-d-1} \dots u_k \quad v_1 v_2 \dots v_k$$

- Caso 3:

$$x = u_{n+1} u_{n+2} \dots u_k \quad v_1 v_2 \dots v_{d+n+1}$$

  con $n \in \{1, 2, \dots, k - d - 2\}$.

- Caso 4:

$$x = u_{k-d-1+n} u_{k-d+n} \dots u_k \quad v_1 v_2 \dots v_k \quad w_1 w_2 \dots w_{d+1-n}$$

  con $n \in \{1, 2, \dots, d\}$

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

00000000 (00000001 0000)0010 00000011 00000100

- Caso 1:

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

- Caso 2:

$$x = u_{k-d-1} \dots u_k \quad v_1 v_2 \dots v_k$$

- Caso 3:

$$x = u_{n+1} u_{n+2} \dots u_k \quad v_1 v_2 \dots v_{d+n+1}$$

  con $n \in \{1, 2, \dots, k - d - 2\}$.

- Caso 4:

$$x = u_{k-d-1+n} u_{k-d+n} \dots u_k \quad v_1 v_2 \dots v_k \quad w_1 w_2 \dots w_{d+1-n}$$

  con $n \in \{1, 2, \dots, d\}$

# Que hay adentro de $X(k)$

Recordemos que definimos $k = 2^d$. Veamos como son las palabras de tamaño $d + k + 1$ que suceden adentro de $X(k)$. Tomemos a modo de ejemplo $k = 8$ y $d = 3$.

$$11111011 \qquad 11111100 \qquad 11111101 \qquad 11111(110 \qquad 11111111 \qquad 0\ldots$$

- Caso 1:

$$x = u_1 u_2 \ldots u_k \quad v_1 v_2 \ldots v_d v_{d+1}$$

- Caso 2:

$$x = u_{k-d-1} \ldots u_k \quad v_1 v_2 \ldots v_k$$

- Caso 3:

$$x = u_{n+1} u_{n+2} \ldots u_k \quad v_1 v_2 \ldots v_{d+n+1}$$

  con $n \in \{1, 2, \ldots, k - d - 2\}$.

- Caso 4:

$$x = u_{k-d-1+n} u_{k-d+n} \ldots u_k \quad v_1 v_2 \ldots v_k \quad w_1 w_2 \ldots w_{d+1-n}$$

  con $n \in \{1, 2, \ldots, d\}$

- Caso 5: Si $x$ comienza al final de $X(k)$ o cerca, se pueden llegar a necesitar palabras fuera de $X(k)$ para completar $d + k + 1$.

# La función $next(w)$

Definimos la siguiente función que nos va a ser útil.

## Definición

La función $next(w) : A^n \to A^n$ se define de la siguiente manera. Si $w$ es la palabra de $n$ 1s, $next(w)$ es la palabra de $n$ 0s. Si no, $next(w)$ es la palabra siguiente en orden lexicográfico.

Por ejemplo,

$$next(0000) = 0001$$

$$next(0001) = 0010$$

$$\vdots$$

$$next(1110) = 1111$$

$$next(1111) = 0000$$

# Análisis por caso - Caso 1

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

Este caso representa a la ocurrencia alineada de $u$ de tamaño $k$ seguido por los primeros $d+1$ símbolos de $v$

$$
\begin{array}{ll}
(00000000 & 0000)\ 0001 \\
(00000001 & 0000)\ 0010 \\
(00000010 & 0000)\ 0011 \\
 \vdots & \\
(00001110 & 0000)\ 1111 \\
(00001111 & 0001)\ 0000 \\
(00010000 & 0001)\ 0001 \\
 \vdots & \\
(11111110 & 1111)\ 1111
\end{array}
$$

# Análisis por caso - Caso 1

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

Este caso representa a la ocurrencia alineada de $u$ de tamaño $k$ seguido por los primeros $d+1$ símbolos de $v$

$(00000000 \qquad 0000)\ 0001$

$(00000001 \qquad 0000)\ 0010$

$(00000010 \qquad 0000)\ 0011$

$\vdots$

$(00001110 \qquad 0000)\ 1111$

$(00001111 \qquad 0001)\ 0000$

$(00010000 \qquad 0001)\ 0001$

$\vdots$

$(11111110 \qquad 1111)\ 1111$

# Análisis por caso - Caso 1

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

Este caso representa a la ocurrencia alineada de $u$ de tamaño $k$ seguido por los primeros $d+1$ símbolos de $v$

$$(00000000 \qquad 0000)\ 0001$$
$$(00000001 \qquad 0000)\ 0010$$
$$(00000010 \qquad 0000)\ 0011$$
$$\vdots$$
$$(00001110 \qquad 0000)\ 1111$$
$$(00001111 \qquad 0001)\ 0000$$
$$(00010000 \qquad 0001)\ 0001$$
$$\vdots$$
$$(11111110 \qquad 1111)\ 1111$$

$$\underbrace{A}_{d+1} \qquad \underbrace{B}_{k-d-1} \qquad A$$

$$\underbrace{A}_{d+1} \qquad \underbrace{11\dots1}_{k-d-1} \qquad next(A)$$

# Análisis por caso - Caso 1

$$x = u_1 u_2 \dots u_k \quad v_1 v_2 \dots v_d v_{d+1}$$

Este caso representa a la ocurrencia alineada de $u$ de tamaño $k$ seguido por los primeros $d+1$ símbolos de $v$

$$
\begin{array}{lll}
(00000000 & 0000) & 0001 \\
(00000001 & 0000) & 0010 \\
(00000010 & 0000) & 0011 \\
 & \vdots & \\
(00001110 & 0000) & 1111 \\
(00001111 & 0001) & 0000 \\
(00010000 & 0001) & 0001 \\
 & \vdots & \\
(11111110 & 1111) & 1111
\end{array}
$$

$$\underbrace{A}_{d+1} \quad \underbrace{B}_{k-d-1} \quad A$$

$$\underbrace{A}_{d+1} \quad \underbrace{11\dots1}_{k-d-1} \quad next(A)$$

Del primer esquema tenemos:

$$2^{d+1}(2^{k-d-1} - 1) = 2^k - \frac{1}{2 \cdot 2^d}$$

Mientras que del segundo:

$$2^{d+1} - 1$$

Juntos son menos que $2^k - 2 \cdot 2^d$

# Análisis por caso - Caso 2

$$x = u_{k-d-1} \ldots u_k \quad v_1 v_2 \ldots v_k$$

Este caso representa a la ocurrencia alineada de $v$ de tamaño $k$ precedida por los últimos $d+1$ símbolos de $u$

$$0000\ (0000 \qquad 00000001)$$

$$0000\ (\ 0001 \qquad 00000010)$$

$$\vdots$$

$$0000\ (1111 \qquad 10000000)$$

$$1000\ (0000 \qquad 10000001)$$

$$\vdots$$

$$1111\ (1110 \qquad 11111111)$$

# Análisis por caso - Caso 2

$$x = u_{k-d-1} \dots u_k \quad v_1 v_2 \dots v_k$$

Este caso representa a la ocurrencia alineada de $v$ de tamaño $k$ precedida por los últimos $d+1$ símbolos de $u$

$0000\ (0000 \qquad 00000001)$

$0000\ (\ 0001 \qquad 00000010)$

$\vdots$

$0000\ (1111 \qquad 10000000)$

$1000\ (0000 \qquad 10000001)$

$\vdots$

$1111\ (1110 \qquad 11111111)$

# Análisis por caso - Caso 2

$$x = u_{k-d-1} \dots u_k \quad v_1 v_2 \dots v_k$$

Este caso representa a la ocurrencia alineada de $v$ de tamaño $k$ precedida por los últimos $d+1$ símbolos de $u$

$0000\ (0000 \quad 00000001)$

$0000\ (\ 0001 \quad 00000010)$

$\vdots$

$0000\ (1111 \quad 10000000)$

$1000\ (0000 \quad 10000001)$

$\vdots$

$1111\ (1110 \quad 11111111)$

$$\underbrace{A}_{d+1} \qquad \underbrace{B}_{k-d-1} \qquad next(A)$$

# Análisis por caso - Caso 2

$$x = u_{k-d-1} \dots u_k \quad v_1 v_2 \dots v_k$$

Este caso representa a la ocurrencia alineada de $v$ de tamaño $k$ precedida por los últimos $d+1$ símbolos de $u$

$0000\ ({\color{red}0000} \qquad 0000{\color{blue}0001})$

$0000\ ({\color{red}0001} \qquad 0000{\color{blue}0010})$

$\vdots$

$0000\ ({\color{red}1111} \qquad 1000{\color{blue}0000})$

$1000\ ({\color{red}0000} \qquad 1000{\color{blue}0001})$

$\vdots$

$1111\ ({\color{red}1110} \qquad 1111{\color{blue}1111})$

$$\underbrace{{\color{red}A}}_{d+1} \qquad \underbrace{B}_{k-d-1} \qquad {\color{blue}next(A)}$$

Este esquema nos da:

$$(2^{d+1}2^{k-d}) - 1 = 2 \cdot 2^k - 1$$

palabras distintas.
Que es menos que $2 \cdot 2^k$ palabras distintas.