

Tesis de Licenciatura en Ciencias de la Computación

# Números Muy Normales

Lucas Puterman

Directora: Verónica Becher  
Codirector: Olivier Carton  
Departamento de Computación  
Facultad de Ciencias Exactas y Naturales  
Universidad de Buenos Aires  
19 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?

► 111111111111111111111111111111111111...

# 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?

- ▶ 111111111111111111111111111111111111...
- ▶ 010101010101010101010101010101010101...

# 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?

- ▶ 111111111111111111111111111111111111...
- ▶ 010101010101010101010101010101010101...
- ▶ 10000110001010001110010010110011010...

## 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.

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.

## Definición (Borel, 1909)

Dado un alfabeto  $A$  y alguna secuencia infinita  $u \in A^\omega$ , decimos que  $u$  es *normal* si para todos los tamaños de bloque  $\ell$ , sucede que todos los bloques posibles de tamaño  $\ell$  ocurren con la misma frecuencia.

# Secuencias normales

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

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.

## Definición (Borel, 1909)

Dado un alfabeto  $A$  y alguna secuencia infinita  $u \in A^\omega$ , decimos que  $u$  es *normal* si para todos los tamaños de bloque  $\ell$ , sucede que todos los bloques posibles de tamaño  $\ell$  ocurren con la misma frecuencia.

## 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

Teorema (Champernowne, 1933)

*La secuencia*

1234567891011121314151617181920...

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

## *champ*, La secuencia que usaremos

### Teorema (Champernowne, 1933)

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

## *champ*, La secuencia que usaremos

### Teorema (Champernowne, 1933)

*Sea  $A$  un alfabeto. Llamamos  $X(n)$  a la concatenación de todas las palabras de longitud  $n$  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:

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

## *champ*, La secuencia que usaremos

### Teorema (Champernowne, 1933)

*Sea  $A$  un alfabeto. Llamamos  $X(n)$  a la concatenación de todas las palabras de longitud  $n$  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

Zeev Rudnick de la Universidad de Tel Aviv definió hace unos años la noción de supernormalidad con la que trabajamos en esta tesis.

Benjamin Weiss de la Universidad Hebrea de Jerusalem dio el 16 de Junio de 2010 una conferencia en el Instituto de Altos Estudios de Princeton titulada “Random-like behavior in deterministic systems” donde describe la noción de supernormalidad, a la que llama “Poisson generic”. Lo poco que se conoce sobre esta noción no está publicado.



Figura: Benjamin Weiss y Zeev Rudnick

## Que tal si contamos

$$x = 10011110 \dots$$

Contemos cuales son las palabras de tamaño 3 que aparecen en los primeros 8 símbolos de  $x$ .

## Que tal si contamos

$$x = 10011110 \dots$$

Contemos cuales son las palabras de tamaño 3 que aparecen en los primeros 8 símbolos de  $x$ .

100

## Que tal si contamos

$$x = 10011110 \dots$$

Contemos cuales son las palabras de tamaño 3 que aparecen en los primeros 8 símbolos de  $x$ .

100, 001



## Que tal si contamos

$$x = 10011110 \dots$$

Contemos cuales son las palabras de tamaño 3 que aparecen en los primeros 8 símbolos de  $x$ .

100, 001, 011

## Que tal si contamos

$$x = 10011110 \dots$$

Contemos cuales son las palabras de tamaño 3 que aparecen en los primeros 8 símbolos de  $x$ .

100, 001, 011, 111

## Que tal si contamos

$$x = 10011110 \dots$$

Contemos cuales son las palabras de tamaño 3 que aparecen en los primeros 8 símbolos de  $x$ .

100, 001, 011, 111, 111

## Que tal si contamos

$$x = 10011110 \dots$$

Contemos cuales son las palabras de tamaño 3 que aparecen en los primeros 8 símbolos de  $x$ .

100, 001, 011, 111, 111, 110

## Que tal si contamos

$$x = 10011110 \dots$$

Contemos cuales son las palabras de tamaño 3 que aparecen en los primeros 8 símbolos de  $x$ .

100, 001, 011, 111, 111, 110

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

Palabra	Cantidad
000	0
001	1
010	0
011	1
100	1
101	0
110	1
111	2

## Que tal si contamos

$$x = 10011110 \dots$$

Contemos cuales son las palabras de tamaño 3 que aparecen en los primeros 8 símbolos de  $x$ .

100, 001, 011, 111, 111, 110

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

Palabra	Cantidad
000	0
001	1
010	0
011	1
100	1
101	0
110	1
111	2

$k$	Cant	Frec
0	3	$3/8$
1	4	$1/2$
2	1	$1/8$
3	0	0
4	0	0
5	0	0
6	0	0
7	0	0
8	0	0

# 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:

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

# 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:

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

## 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

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

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



El resultado de esta tesis

# 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  $\nRightarrow$  supernormal )*
- 2. Sea  $x$  una secuencia infinita. Si  $x$  es supernormal, entonces  $x$  es normal. (supernormal  $\Rightarrow$  normal )*

# 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  $\nRightarrow$  supernormal )*
2. *Sea  $x$  una secuencia infinita. Si  $x$  es supernormal, entonces  $x$  es normal. (supernormal  $\Rightarrow$  normal )*

"Who are you?"

"I'm you, but stronger"



## ¿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 orden lexicográfico.

Llamamos *champ* a la concatenación de  $X(n)$  para  $n = 1, 2, \dots$

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

## ¿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  $n = 1, 2, \dots$

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

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

## ¿Es *champ* supnormal?

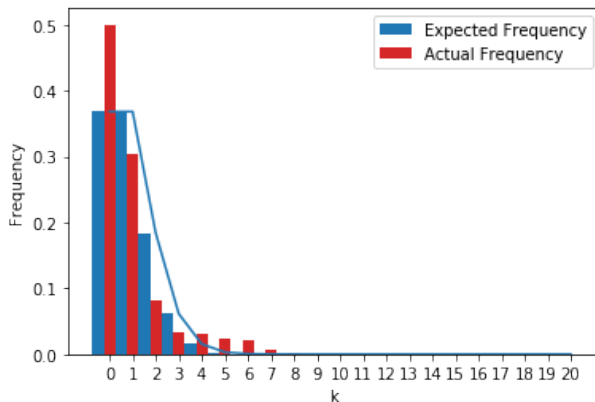


Figura: Frecuencias observadas y esperadas en *champ* for  $n = 16$  y  $\lambda = 1$ .

## ¿Es *champ* supernormal?

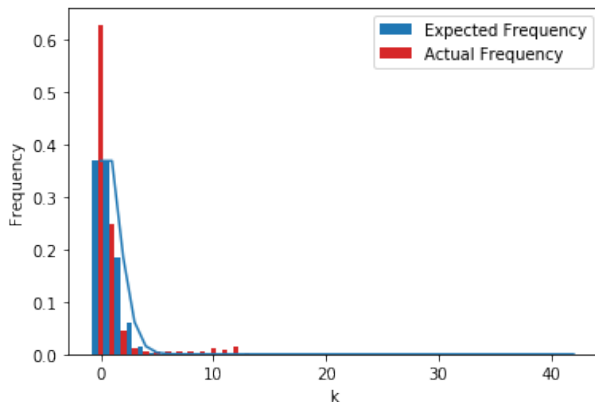


Figura: Frecuencias observadas y esperadas en *champ* for  $n = 22$  y  $\lambda = 1$ .

## Normal $\nRightarrow$ Supernormal - Idea de la demostración

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



## Normal $\nRightarrow$ Supernormal - Idea de la demostración

- ▶ Sabemos que  $champ = X(1)X(2)X(3)\dots$
- ▶ Vamos a buscar el  $k$  más grande tal que  $X(k)$  este completamente contenido en  $champ[1 \dots 2^n]$  y contar cuales son las palabras que aparecen adentro.

## Normal $\nRightarrow$ Supernormal - Idea de la demostración

- ▶ Sabemos que  $champ = X(1)X(2)X(3)\dots$
- ▶ Vamos a buscar el  $k$  más grande tal que  $X(k)$  este completamente contenido en  $champ[1 \dots 2^n]$  y contar cuales son las palabras que aparecen adentro.
- ▶ Para esto buscamos una forma de elegir algunos  $ns$  que tengan una pinta que nos sirva.

# Normal $\nRightarrow$ Supernormal - Idea de la demostración

- ▶ Sabemos que  $champ = X(1)X(2)X(3)\dots$
- ▶ Vamos a buscar el  $k$  más grande tal que  $X(k)$  este completamente contenido en  $champ[1\dots 2^n]$  y contar cuales son las palabras que aparecen adentro.
- ▶ Para esto buscamos una forma de elegir algunos  $ns$  que tengan una pinta que nos sirva.
- ▶ Encontramos que si tomamos  $n = d + k + 1$  con  $k = 2^d$ , sucede que  $X(k)$  está completamente contenido en  $champ[1\dots 2^n]$  y además ocupa la mitad del espacio.

## Normal $\nRightarrow$ Supernormal - Idea de la demostración

- ▶ Sabemos que  $champ = X(1)X(2)X(3)\dots$
- ▶ Vamos a buscar el  $k$  más grande tal que  $X(k)$  este completamente contenido en  $champ[1\dots 2^n]$  y contar cuales son las palabras que aparecen adentro.
- ▶ Para esto buscamos una forma de elegir algunos  $n$ s que tengan una pinta que nos sirva.
- ▶ Encontramos que si tomamos  $n = d + k + 1$  con  $k = 2^d$ , sucede que  $X(k)$  está completamente contenido en  $champ[1\dots 2^n]$  y además ocupa la mitad del espacio.
- ▶ Nos fijamos cuales son las palabras de longitud  $d + k + 1$  que aparecen adentro de  $X(k)$ .

## Normal $\nRightarrow$ Supernormal - Idea de la demostración

- ▶ Sabemos que  $champ = X(1)X(2)X(3)\dots$
- ▶ Vamos a buscar el  $k$  más grande tal que  $X(k)$  este completamente contenido en  $champ[1\dots 2^n]$  y contar cuales son las palabras que aparecen adentro.
- ▶ Para esto buscamos una forma de elegir algunos  $ns$  que tengan una pinta que nos sirva.
- ▶ Encontramos que si tomamos  $n = d + k + 1$  con  $k = 2^d$ , sucede que  $X(k)$  está completamente contenido en  $champ[1\dots 2^n]$  y además ocupa la mitad del espacio.
- ▶ Nos fijamos cuales son las palabras de longitud  $d + k + 1$  que aparecen adentro de  $X(k)$ .
- ▶ Conociendo las palabras que aparecen, vemos que las son más 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

00000001

00000010

00000011

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    0000)0001    00000010    00000011    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 00000)001      00000010      00000011      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  $d = 3$ .

00(000000 000000)01      00000010      00000011      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  $d = 3$ .

000(00000      0000000)1      00000010      00000011      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      00000001)      00000010      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  $n \in \{1, 2, \dots, k - d - 2\}$ .

## 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      00000001      0)0000010      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  $d = 3$ .

000000(00      00000001      00)000010      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  $d = 3$ .

0000000(0      00000001      000)00010      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  $d = 3$ .

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      11111100      11111101      11111(110      11111111      0...

► 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  $n \in \{1, 2, \dots, d\}$

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

# 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      0000) 0001

(00000001      0000) 0010

(00000010      0000) 0011

⋮

(00001110      0000) 1111

(00001111      0001) 0000

(00010000      0001) 0001

⋮

(11111110      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 0000) 0001

(00000001 0000) 0010

(00000010 0000) 0011

⋮

(00001110 0000) 1111

(00001111 0001) 0000

(00010000 0001) 0001

⋮

(11111110 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 0000) 0001

(00000001 0000) 0010

(00000010 0000) 0011

⋮

(00001110 0000) 1111

(00001111 0001) 0000

(00010000 0001) 0001

⋮

(11111110 1111) 1111

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

# 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  $v$

(00000000 0000) 0001

(00000001 0000) 0010

(00000010 0000) 0011

⋮

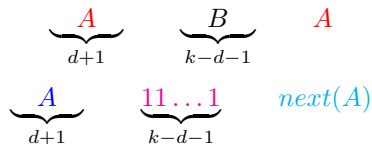
(00001110 0000) 1111

(00001111 0001) 0000

(00010000 0001) 0001

⋮

(11111110 1111) 1111



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

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

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

0000 (0000      00000001)

0000 ( 0001      00000010)

⋮

0000 (1111      10000000)

1000 (0000      10000001)

⋮

1111 (1110      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 00000001)

0000 ( 0001 00000010)

⋮

0000 (1111 10000000)

1000 (0000 10000001)

⋮

1111 (1110 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 00000001)

0000 ( 0001 00000010)

⋮

0000 (1111 10000000)

1000 (0000 10000001)

⋮

1111 (1110 11111111)

$\underbrace{A}_{d+1}$

$\underbrace{B}_{k-d-1}$

$next(A)$



## 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 (0000 00000001)

0000 ( 0001 00000010)

⋮

0000 (1111 10000000)

1000 (0000 10000001)

⋮

1111 (1110 11111111)

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

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.

## Análisis por caso - Caso 3

$$x = u_{n+1}u_{n+2} \dots u_k \quad v_1v_2 \dots v_{d+n+1} \quad n \in \{1, 2, \dots, k - d - 2\}$$

Este caso representa cuando los  $k + d + 1$  símbolos son tomados de dos palabras  $u$  y  $v$  de longitud  $k$  y ninguna de las dos está completas.

Ejemplo tomando  $n = 3$ ,  $k = 8$  y  $d = 3$ .

000 (0000 0      000 0000) 1

000 (0000 1      000 0001) 0

000 (0001 0      000 0001) 1

000 (0001 1      000 0010) 0

⋮

111 (1110 1      111 1111) 0

111 (1111 0      111 1111) 1

## Análisis por caso - Caso 3

$$x = u_{n+1}u_{n+2} \dots u_k \quad v_1v_2 \dots v_{d+n+1} \quad n \in \{1, 2, \dots, k - d - 2\}$$

Este caso representa cuando los  $k + d + 1$  símbolos son tomados de dos palabras  $u$  y  $v$  de longitud  $k$  y ninguna de las dos está completas.

Ejemplo tomando  $n = 1$ ,  $k = 8$  y  $d = 3$ .

0 (0000 000      0 0000) 001

0 (0000 001      0 0000) 010

⋮

0 (0001 110      0 0001) 111

0 (0001 111      0 0010) 000

0 (0010 000      0 0010) 001

⋮

1 (1111 101      1 1111) 110

0 (1111 110      1 1111) 111

## Análisis por caso - Caso 3

$$x = u_{n+1}u_{n+2} \dots u_k \quad v_1v_2 \dots v_{d+n+1} \quad n \in \{1, 2, \dots, k - d - 2\}$$

Este caso representa cuando los  $k + d + 1$  símbolos son tomados de dos palabras  $u$  y  $v$  de longitud  $k$  y ninguna de las dos está completas.

Ejemplo tomando  $n = 1$ ,  $k = 8$  y  $d = 3$ .

0 (0000 000      0 0000) 001

0 (0000 001      0 0000) 010

⋮

0 (0001 110      0 0001) 111

0 (0001 111      0 0010) 000

0 (0010 000      0 0010) 001

⋮

1 (1111 101      1 1111) 110

0 (1111 110      1 1111) 111

# Análisis por caso - Caso 3

$$x = u_{n+1}u_{n+2} \dots u_k \quad v_1v_2 \dots v_{d+n+1} \quad n \in \{1, 2, \dots, k - d - 2\}$$

Este caso representa cuando los  $k + d + 1$  símbolos son tomados de dos palabras  $u$  y  $v$  de longitud  $k$  y ninguna de las dos está completa.

Ejemplo tomando  $n = 1$ ,  $k = 8$  y  $d = 3$ .

0 (0000 000      0 0000) 001

0 (0000 001      0 0000) 010

⋮

0 (0001 110      0 0001) 111

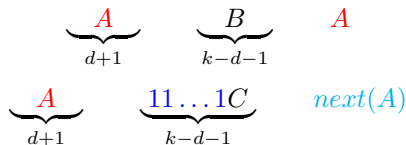
0 (0001 111      0 0010) 000

0 (0010 000      0 0010) 001

⋮

1 (1111 101      1 1111) 110

0 (1111 110      1 1111) 111



# Análisis por caso - Caso 3

$$x = u_{n+1}u_{n+2} \dots u_k \quad v_1v_2 \dots v_{d+n+1} \quad n \in \{1, 2, \dots, k - d - 2\}$$

Este caso representa cuando los  $k + d + 1$  símbolos son tomados de dos palabras  $u$  y  $v$  de longitud  $k$  y ninguna de las dos está completa.

Ejemplo tomando  $n = 1$ ,  $k = 8$  y  $d = 3$ .

0 (0000 000      0 0000) 001

0 (0000 001      0 0000) 010

⋮

0 (0001 110      0 0001) 111

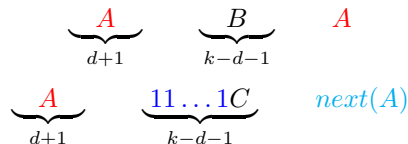
0 (0001 111      0 0010) 000

0 (0010 000      0 0010) 001

⋮

1 (1111 101      1 1111) 110

0 (1111 110      1 1111) 111



Pero el primer esquema ya lo contamos en el caso 1 y el segundo esquema es un caso particular de uno de los esquemas del caso 2. Por lo que el caso 3 no nos arroja nuevas palabras.

## Análisis por caso - Caso 4

$$x = u_{k-d-1+n}u_{k-d+n} \dots u_k \quad v_1v_2 \dots v_k \quad w_1w_2 \dots w_{d+1-n}$$

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

Este caso representa cuando los  $k + d + 1$  símbolos son tomados de tres palabras consecutivas  $u$ ,  $v$  y  $w$  de longitud  $k$  de los cuales  $k$  símbolos son tomados de  $v$  y los restantes de  $u$  y  $w$ . Ejemplo tomando  $n = 1$ ,  $k = 8$  y  $d = 3$ .

0000000 (0      000 0000 1      000)00010

0000000 (1      000 0001 0      000)00011

⋮

0011111 (0      001 1111 1      001)10000

0011111 (1      010 0000 0      010)00001

⋮

1111110 (1      111 1111 0      111)11111

## Análisis por caso - Caso 4

$$x = u_{k-d-1+n}u_{k-d+n} \dots u_k \quad v_1v_2 \dots v_k \quad w_1w_2 \dots w_{d+1-n}$$

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

Este caso representa cuando los  $k + d + 1$  símbolos son tomados de tres palabras consecutivas  $u$ ,  $v$  y  $w$  de longitud  $k$  de los cuales  $k$  símbolos son tomados de  $v$  y los restantes de  $u$  y  $w$ . Ejemplo tomando  $\mathbf{n = 2}$ ,  $k = 8$  y  $d = 3$ .

000000 (00      00 0000 01      00)000010

000000 (01      00 0000 10      00)000011

⋮

001111 (01      00 1111 10      00)111111

001111 (10      00 1111 11      01)000000

001111 (11      01 0000 00      01)000001

⋮

111111 (01      11 1111 10      11)111111



## Análisis por caso - Caso 4

$$x = u_{k-d-1+n}u_{k-d+n} \dots u_k \quad v_1v_2 \dots v_k \quad w_1w_2 \dots w_{d+1-n}$$

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

000000 (00 00 0000 01 00)000010

000000 (01 00 0000 10 00)000011

$\vdots$

001111 (01 00 1111 10 00)111111

001111 (10 00 1111 11 01)000000

001111 (11 01 0000 00 01)000001

$\vdots$

111111 (01 11 1111 10 11)111111

# Análisis por caso - 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  $n \in \{1, 2, \dots, d\}$

000000 (00 00 0000 01 00)000010

000000 (01 00 0000 10 00)000011

⋮

001111 (01 00 1111 10 00)111111

001111 (10 00 1111 11 01)000000

001111 (11 01 0000 00 01)000001

⋮

111111 (01 11 1111 10 11)111111

$$\underbrace{A}_n \quad \underbrace{B}_{d+1-n} \quad \underbrace{C}_{k-d+1} \quad \text{next}(A) \quad B$$

$$\underbrace{1\dots 10}_n \quad \underbrace{B}_{d+1-n} \quad \underbrace{1\dots 1}_{k-d+1} \quad n(A) \quad n(B)$$

Para el primer esquema tenemos

$$2^k \sum_{n=1}^d \frac{2^n - 1}{2^n}$$

Para el segundo esquema  $2 \cdot 2^d \sum_{n=1}^d 2^{-n}$

Juntos son menos que

$$d2^k + 2 \cdot 2^d$$

## Hay demasiadas palabras que no aparecen en *champ*

Si *champ* fuera supernormal, la frecuencia esperada de palabras que aparecen una o más veces en los primeros  $2^n$  símbolos debería satisfacer:

$$\lim_{n \rightarrow \infty} \frac{\#\{w : |w| = n, |champ[1.., 2^n]|_w > 0\}}{2^n} = 1 - e^{-1}$$

## Hay demasiadas palabras que no aparecen en *champ*

Si *champ* fuera supernormal, la frecuencia esperada de palabras que aparecen una o más veces en los primeros  $2^n$  símbolos debería satisfacer:

$$\lim_{n \rightarrow \infty} \frac{\#\{w : |w| = n, |champ[1.., 2^n]|_w > 0\}}{2^n} = 1 - e^{-1}$$

Sabemos que  $X(k)$  ocupa la mitad de las palabras de los primeros  $2^{d+k+1}$  símbolos de *champ*. Si analizamos lo que sucede en las palabras de longitud  $d+k+1$  usando las cotas que calculamos en los diferentes casos, y asumimos que la mitad restante ( $2^{d+k} + d + k - 1$ ) son todas diferentes, entonces:

$$\lim_{d \rightarrow \infty} \frac{\#\{w : |w| = d + k + 1, |champ[1 \dots 2^{d+k+1}]|_w > 0\}}{2^{d+k+1}} <$$
$$\lim_{d \rightarrow \infty} \frac{\text{Caso 1} + \text{Caso 2} + \text{Caso 4} + \text{Caso 5} + \text{otra mitad}}{2^{d+k+1}} =$$

## Hay demasiadas palabras que no aparecen en *champ*

Si *champ* fuera supernormal, la frecuencia esperada de palabras que aparecen una o más veces en los primeros  $2^n$  símbolos debería satisfacer:

$$\lim_{n \rightarrow \infty} \frac{\#\{w : |w| = n, |champ[1.., 2^n]|_w > 0\}}{2^n} = 1 - e^{-1}$$

Sabemos que  $X(k)$  ocupa la mitad de las palabras de los primeros  $2^{d+k+1}$  símbolos de *champ*. Si analizamos lo que sucede en las palabras de longitud  $d+k+1$  usando las cotas que calculamos en los diferentes casos, y asumimos que la mitad restante ( $2^{d+k} + d + k - 1$ ) son todas diferentes, entonces:

$$\begin{aligned} \lim_{d \rightarrow \infty} \frac{\#\{w : |w| = d+k+1, |champ[1 \dots 2^{d+k+1}]|_w > 0\}}{2^{d+k+1}} &< \\ \lim_{d \rightarrow \infty} \frac{\text{Caso 1} + \text{Caso 2} + \text{Caso 4} + \text{Caso 5} + \text{otra mitad}}{2^{d+k+1}} &= \\ \lim_{d \rightarrow \infty} \frac{(2^k - 2 \cdot 2^d) + (2 \cdot 2^k) + (d2^k + 2 \cdot 2^d) + (2d + k + 1) + (2^{d+k})}{2^{d+k+1}} &= \\ \frac{1}{2} &< 1 - e^{-1} \end{aligned}$$

## Acotando palabras distintas en *champ*

Finalmente, si en *champ* consideramos que los  $n$ 's tales que  $n = d + 2^d + 1$  son una subsecuencia de  $n = 1, 2, \dots$ , podemos afirmar que si:

$$\lim_{n \rightarrow \infty} \frac{\#\{w : |w| = n, |champ[1.., 2^n]|_w > 0\}}{2^n}$$

existe, este no es  $1 - e^{-1}$ .

### Corolario

*Si  $x$  es normal, no necesariamente  $x$  es supernormal.*

# Agradecimiento y Trabajo futuro

Esta tesis fue realizada con la dirección conjunta de Verónica Becher y Olivier Carton en el marco del Laboratoire International Associé SINFIN, Université Paris Diderot-CNRS/Universidad de Buenos Aires-CONICET)

- ▶ Demostrar que casi todas las palabras son supernormales
- ▶ Encontrar una palabra supernormal (Dar un algoritmo)
- ▶ Caracterizar la supernormalidad en términos de compresión

¡Fin!

¿Preguntas?





BAILEY, D. H., AND MISUREWICZ, M.

**A strong hot spot theorem.**

*Proceedings of the American Mathematical Society* 134, 9 (2006), 2495–2501.



BECHER, V., AND CARTON, O.

**Normal numbers and computer science.**

In *Sequences, Groups, and Number Theory*, V. Berthé and M. Rigo, Eds. Springer International Publishing, Cham, 2018, pp. 233–269.



BECHER, V., CARTON, O., AND CUNNINGHAM, I.

**Low discrepancy sequences failing poissonian pair correlations.**

*Archiv der Mathematik* 113 (04 2019).



BOREL, É.

**Les probabilités dénombrables et leurs applications arithmétiques.**

*Rendiconti del Circolo Matematico di Palermo (1884-1940)* 27, 1 (Dec 1909), 247–271.



BUGEAUD, Y.

***Distribution Modulo One and Diophantine Approximation.***

No. 193 in Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, UK, 2012.



CHAMPERNOWNE, D. G.

**The Construction of Decimals Normal in the Scale of Ten.**

*Journal of the London Mathematical Society* s1-8, 4 (10 1933), 254–260.



FIGUIERA, S., AND NIES, A.

**Feasible analysis and randomness.**

*Theory Comput Syst* (2015) 56 (2015), 439.



KUIPERS, L., AND NIEDERREITER, H.

***Uniform distribution of sequences.***

Dover Publications, Inc., New York, 2006.



PIRSIC, I., AND STOCKINGER, W.

**The champernowne constant is not poissonian.**

*Funct. Approx. Comment. Math.* 60, 2 (03 2019), 253–262.



SIERPINSKI, W.

**Démonstration élémentaire du théorème de m. borel sur les nombres absolument normaux et détermination effective d'une tel nombre.**

*Bulletin de la Société Mathématique de France* 45 (1917), 125–132.



WEISS, B.

Random-like behavior in deterministic systems - <https://youtu.be/8ab7591de68?t=1280>, 2010.