Tesis de Licenciatura

Sobre el trabajo no publicado de Alan M.Turing "A note on normal numbers"

Rafael Eduardo Picchi

LU:647/92 - *rp71@dc.uba.ar*

Directora: Dra. Verónica Becher

vbecher@dc.uba.ar

Codirector: Lic. Santiago Figueira

sfigueir@dc.uba.ar

Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales
Departamento de Computación

Buenos Aires - Argentina

Diciembre de 2005

A mi viejo,
y
a Cecilia.

Agradecimientos:

A Verónica Becher, mi directora de tesis, por su infinita predisposición, su optimismo constante y su inagotable aliento.

A mi codirector, Santiago Figueira, por su enorme voluntad y su valiosa ayuda en las distintas etapas de este trabajo.

A Mariela Sued, por su asesoramiento y colaboración, que se ve reflejado en la sección 5.

A Max Dickmann y a Pablo Jacovkis, por sus participaciones en distintas discusiones que nos permitieron visualizar otros enfoques.

A mis hermanos, Conrado y Silvana, por sus aportes desde los primeros pasos de este trabajo.

A Graciela, mi compañera en la vida, por su apoyo incondicional, sin el cual me hubiera sido imposible realizar esta tesis.

Resumen

Esta tesis está basada en el manuscrito no publicado de Alan Turing titulado "A note on normal numbers". El manuscrito de Turing tiene tiene por objetivo dar dos teoremas. El primero es una demostración constructiva de que la mayoría (en el sentido de la medida de Lebesgue) de los números reales son absolutamente normales. Este teorema fue probado con anterioridad por Borel en 1909, pero de una manera no constructiva. El segundo teorema es un algoritmo para generar instancias de números absolutamente normales. En el manuscrito de Turing ninguna de las dos demostraciones están completamente desarrolladas. En esta tesis damos una reconstrución completa del Teorema 1 de Turing.

El interés de este trabajo es conocer las técnicas que utilizó Turing en relación a los números normales, especialmente porque actualmente no se cuenta con métodos que permitan demostrar la normalidad de números reales, ni se conocen algoritmos rápidos para dar instancias de números normales.

Índice

1. Introducción **2**

2. La definición de normalidad **4**

3. Reconstrucción del Teorema 1 de Turing **5**

3.1. Comparando las cotas 13

4. Versión fiel del Teorema 1 de Turing **15**

5. Una cota alternativa **19**

6. Conclusiones **22**

7. Apéndice 1
Manuscritos Originales de A.M.Turing **24**

8. Apéndice 2
Versión transcripta por J.L.Britton **43**

1. Introducción

En este trabajo hacemos una revisión del manuscrito de Alan M. Turing titulado "A note on normal numbers", trabajo que ha permanecido inédito hasta su reciente aparición en las obras completas de Alan Turing en la colección "Collected Works of A.M.Turing", editada por J.L.Britton (1992) [11], pp. 117-119, con notas del editor en pp. 263-265 [1]. Britton destaca las dificultades para la transcripción del manuscrito de Turing debido a que los originales son bastante ilegibles y están incompletos.

Hay una versión "scaneada"del manuscrito original de Turing que puede verse en la dirección de internet `http://www.turingarchive.org` [2].

Nuestra motivación para abordar este trabajo fue conocer las técnicas que utilizó Turing en relación a los números normales, especialmente porque actualmente no se cuenta con métodos que permitan demostrar la normalidad de números reales. Por ejemplo es aún una conjetura que la constante π es normal en base 10. Tampoco se conocen algoritmos rápidos para construir números absolutamente normales ([3, 4, 5]).

En su manuscrito, Turing enuncia dos resultados, con insuficientes detalles de demostración. El primero, su Teorema 1, es una demostración constructiva efectiva de que que casi todos los números reales son absolutamente normales.

Teorema 1 (Turing). *Podemos dar una función recursiva* $c : \mathbb{N} \times \mathbb{N} \to \mathcal{P}(\mathbb{R})$ *y un* $k_0 \in \mathbb{N}$ *tales que, para* $k \geq k_0$*:* $c(k, n+1) \subsetneqq c(k, n)$ *y*

$$E(k) = \bigcap_{n \geq 1} c(k, n)$$

tiene medida 1 -1/k y contiene únicamente reales absolutamente normales.

El segundo resultado es un método para generar constructivamente números absolutamente normales particulares. Es decir, un algoritmo para producir números absolutamente normales. Este Teorema 2 utiliza el Teorema 1.

Nuestro trabajo ha sido reconstruir el Teorema 1 de Turing y dar todos los detalles de su demostración. Naturalmente nuestro interés en el Teorema 1 es indagar qué herramientas utiliza Turing, ya que el resultado de que los números normales

[1] Observamos que las notas [1] y [2] del editor no son correctas

[2] Este archivo digital contiene principalmente papeles personales no publicados y fotografías de Alan Turing de 1923 a 1972. Los originales se encuentran en el archivo Turing en King's College Cambridge.

tienen medida 1, se debe a Émil Borel y data del año 1909 [2], mediante una demostración no constructiva. Una demostración constructiva pero no recursiva del mismo resultado se debe a Sierpinski [9]. Aunque toman como punto de partida distintas –pero equivalentes– definiciones de normalidad, la demostración de Turing y la de Sierpinski tienen un espíritu similar; y por otro lado, la construcción de Turing tiene cotas más ajustadas que las de Sierpinski.

El trabajo está organizado de la siguiente manera. En la Seccion 2 damos las distintas definiciones equivalentes de normalidad. En la sección 3 hacemos una reconstrucción completa del Teorema 1 de Turing. En la sección 4 desarrollamos la demostración del Teorema 1 de Turing siguiendo fielmente la demostración que figura en el manuscrito. Lamentablemente hay un lema central que no pudimos demostrar. Por último, en la sección 5, damos una cota alternativa a la que utilizamos en nuestra reconstrucción del Teorema 1 de Turing.

En los apéndices 1 y 2, incluiremos las páginas "scaneadas"de los manuscritos originales y la versión transcripta por Britton.

2. La definición de normalidad

Sea t un entero mayor o igual que 2. Una *palabra* en base t de longitud R es una secuencia de R símbolos en el afabeto $\{0, ..., t-1\}$. La longitud de una palabra w la denotamos: $|w|$. Indistintamente, en diferentes partes de nuestro trabajo, utilizaremos el término dígito (en base t) o bien para nombrar a una palabra en base t de longitud 1, o bien, para referenciar a un símbolo del afabeto $\{0, ..., t-1\}$.

Todo número real $\alpha > 0$ tiene una expansión fraccionaria única en base t de la forma

$$\alpha = [\alpha] + \sum_{n=1}^{\infty} a_n t^{-n}$$

donde [ ] denota parte entera, $0 \leq a_n < t$ y $a_n < t-1$ para infinitos valores de n. Esta última condición sobre a_n se introduce para asegurar la representación única de ciertos racionales.

Definición 2. Sea α un número real en $(0,1)$, y γ una palabra en base t. Definimos $S(\alpha, t, \gamma, R)$ como el número de ocurrencias de γ en los primeros R dígitos después del punto decimal, en la expansión de α en base t.

La definición de la propiedad de normalidad para números reales es de Borel, del año 1909 [2].

Definición 3. Sean α un número real en $(0,1)$, $t \in \mathbb{N}, t > 1$, $\forall$ dígito d en base t. α is *simplemente normal* en base t si

$$\lim_{R\to\infty} \frac{S(\alpha, t, d, R)}{R} = \frac{1}{t}$$

α es *absolutamente normal* si es simplemente normal en toda base $t \geq 2$.

Borel indica que los números absolutamente normales tienen una propiedad que los caracteriza: la propiedad de que todos los bloques de igual tamaño aparecen con la misma frecuencia, en cada base posible. Esta propiedad da lugar a la Definición 4 que tiene la apariencia de ser más exigente que la definición de absoluta normalidad dada por la Definicion 3. Las dos definiciones son equivalentes. La demostración de la equivalencia se puede ver en el libro de Harman. [7] (Teorema 1.3, pag 5).

Definición 4. α un número real en $(0,1)$, $t \in \mathbb{N}, t > 1$.
α es *normal* si para cada palabra γ en base t, $\forall t > 1$,

$$\lim_{R\to\infty} \frac{S(\alpha, t, \gamma, R)}{R} = t^{-|\gamma|}$$

En su manuscrito Turing utiliza la Definición 4 de normalidad basada en bloques de dígitos. Nuestra reconstrucción del Teorema 1 utiliza la Definición 3, en el sentido que la construcciones están basadas en dígitos.

3. Reconstrucción del Teorema 1 de Turing

En su demostración del Teorema 1 Turing utiliza una cota superior de la cantidad de palabras de una longitud dada en las que un bloque de dígitos dado tiene demasiadas, o demasiado pocas ocurrencias con respecto a su valor esperado. *Turing no demuestra esta cota* .

Nosotros reformulamos la demostración del Teorema 1 de Turing y evitamos así el resultado faltante. En vez de basarnos en la Definición 4 de normalidad lo hacemos en la Definición 3, por lo que requerimos una cota para la cantidad de palabras de una longitud dada en la que *un dígito* dado aparece en defecto o en exceso respecto de la cantidad de veces esperada. Esta cota se prueba en el Lema 8 de esta sección. Observaremos que la cota obtenida en el Lema 8 es apenas mayor que la utilizada por Turing instanciada para dígitos (bloques de longitud 1).

Definición 5. Sea $t \in \mathbb{N}$, $t > 1$, $n \in \mathbb{N}_0$, $R \in \mathbb{N}$ y γ una palabra en base t. Definimos $N(t, \gamma, n, R)$ como el número de palabras en base t de longitud R, en las que γ ocurre exactamente n veces (incluyendo superposiciones, si las hubiera).

A modo de ejemplo, las siguientes tablas muestran la cantidad de apariciones del dígito 0 y la palabra 11 para cada cadena binaria de longitud 3 y los valores de $N(t = 2, \gamma = 0, n, R = 3)$ y $N(t = 2, \gamma = 11, n, R = 3)$, respectivamente.

Palabra	$\gamma = 0$	$\gamma = 11$
000	3	0
001	2	0
010	2	0
011	1	1
100	2	0
101	1	0
110	1	1
111	0	2

n	$N(2, 0, n, 3)$	$N(2, 11, n, 3)$
0	1	5
1	3	2
2	3	1
3	1	0

Como ya dijimos, Turing da una cota para la cantidad de palabras de una longitud dada en las que un *bloque* de dígitos dado aparece en exceso o en defecto respecto de la cantidad esperada. En particular, la cota de Turing se puede instanciar para bloques de longitud 1, es decir, para *un dígito*. Se puede ver fácilmente, que la cantidad media de apariciones de cualquier dígito d en una palabra de longitud R en base t es R/t. La cota de Turing instanciada para dígitos, es la siguiente:

Misterioso Lema 6 (Turing). *Sea $t \in \mathbb{N}$, $t > 1$ y $\forall d$ dígito en base t. $\forall R \in \mathbb{N}, k \in \mathbb{R}$ tales que: $\frac{kt}{R} < 0.3$:*

$$\sum_{n:|n-R/t|>k} N(t,d,n,R) < 2t^R e^{-\frac{k^2 t}{4R}}$$

Damos ahora el siguiente Lema 8 que usaremos en lugar del no demostrado Misterioso Lema 6. Este Lema surgió de completar los detalles faltantes y expresar convenientemente el Lema 1.1 de Harman [7] .

Observación 7. Por un argumento elemental de combinatoria,
$N(t,d,n,R) = \begin{pmatrix} R \\ n \end{pmatrix} (t-1)^{R-n}$

Lema 8. *Sea $t \in \mathbb{N}$, $t > 1$. Sea d un dígito en base t. Para todo $R \in \mathbb{N}, \varepsilon \in \mathbb{R}$ tales que $\frac{6}{R} \leq \varepsilon \leq \frac{1}{t}$:*

$$\sum_{|n-R/t|>\varepsilon R} N(t,d,n,R) \ < \ 2 \ (1-\varepsilon t/3)^{\lfloor \varepsilon R/2 \rfloor} \, t^R.$$

Demostración.

$$\sum_{\left|n-\frac{R}{t}\right|>\varepsilon R} N(t,d,n,R) \quad \leq \quad \sum_{n=0}^{\lfloor \frac{R}{t}-\varepsilon R \rfloor} N(t,d,n,R) + \sum_{n=\lfloor \frac{R}{t}+\varepsilon R \rfloor}^{R} N(t,d,n,R)$$

Sea $Y = \sum_{n=0}^{\lfloor \frac{R}{t}-\varepsilon R \rfloor} N(t,d,n,R) = \sum_{n=0}^{\lfloor \frac{R}{t}-\varepsilon R \rfloor} \begin{pmatrix} R \\ n \end{pmatrix} (t-1)^{R-n}$.

Claramente, $\forall R : \sum_{n=0}^{R} \begin{pmatrix} R \\ n \end{pmatrix} (t-1)^{R-n} = t^R$.

Sea $a_n = \begin{pmatrix} R \\ n \end{pmatrix} (t-1)^{R-n}$ los términos en la suma Y. Sabemos que para todo n menor que $\frac{R}{t}$, los términos son estrictamente crecientes. Esto es, $a_n < a_{n+1}$ o $\frac{a_n}{a_{n+1}} < 1$.

$$\frac{a_n}{a_{n+1}} = \frac{\binom{R}{n}(t-1)^{R-n}}{\binom{R}{n+1}(t-1)^{R-n-1}} \quad = \quad \frac{(n+1)(t-1)}{R-n}$$

y reemplazando $n = \lfloor \frac{R}{t} - \theta \rfloor$, con $\theta > 0$

$$\begin{aligned}
&= \frac{(t-1)(\lfloor \frac{R}{t} - \theta \rfloor + 1)}{R - \lfloor \frac{R}{t} - \theta \rfloor} \\
&\leq \frac{(t-1)(\frac{R}{t} - \theta + 1)}{R - \frac{R}{t} + \theta} \\
&= 1 - \frac{t\theta - t + 1}{R - \frac{R}{t} + \theta} \\
&= 1 - \frac{t(\theta-1)+1}{R(1-1/t)+\theta}
\end{aligned}$$

para $\theta \leq \frac{R}{t}$

$$< \quad 1 - \frac{t(\theta-1)}{R}$$

para $\theta \geq \varepsilon R/2$

$$\leq \quad 1 - \frac{\varepsilon t}{2} + \frac{t}{R}$$

Usando las hipótesis: $\frac{6}{R} \leq \varepsilon \leq \frac{1}{t}$ tenemos que para cada n tal que $0 \leq n < \frac{R}{t} - \varepsilon R/2$

$$\frac{a_n}{a_{n+1}} \leq \ 1 - \frac{\varepsilon t}{3} \ < 1$$

Damos ahora una cota superior para Y "desplazando la suma hacia la derecha" $\lfloor \varepsilon R/2 \rfloor$ posiciones. Usamos que para cada $0 \leq n \leq \frac{R}{t} - \varepsilon R$,

$$a_n = \frac{a_n}{a_{n+1}} \frac{a_{n+1}}{a_{n+2}} \cdots \frac{a_{n+\lfloor \varepsilon R/2 \rfloor}}{a_{n+\lfloor \varepsilon R/2 \rfloor + 1}} a_{n+\lfloor \varepsilon R/2 \rfloor + 1} < (1 - \frac{\varepsilon t}{3})^{\lfloor \varepsilon R/2 \rfloor} a_{n+\lfloor \varepsilon R/2 \rfloor + 1}$$

$$\begin{aligned}
Y = \sum_{n=0}^{\lfloor \frac{R}{t} - \varepsilon R \rfloor} a_n \quad &< \quad \sum_{n=0}^{\lfloor \frac{R}{t} - \varepsilon R \rfloor} (1 - \frac{\varepsilon t}{3})^{\lfloor \varepsilon R/2 \rfloor} a_{\lfloor \varepsilon R/2 \rfloor + n + 1} \\
&= \quad (1 - \frac{\varepsilon t}{3})^{\lfloor \varepsilon R/2 \rfloor} \sum_{n=0}^{\lfloor \frac{R}{t} - \varepsilon R \rfloor} a_{\lfloor \varepsilon R/2 \rfloor + n + 1} \\
&< \quad (1 - \frac{\varepsilon t}{3})^{\lfloor \varepsilon R/2 \rfloor} t^R \ \ (\textit{porque} \sum_{n=0}^{R} a_n = t^R)
\end{aligned}$$

La misma cota se obtiene para $Z = \sum_{n=\lfloor \varepsilon R + \frac{R}{t} \rfloor}^{R} N(t,d,n,R)$ de la siguente manera. Sean a_n los términos en esa suma. Vamos a ver que $\forall n > \lfloor \frac{R}{t} \rfloor + 1$, $a_n > a_{n+1}$, o sea que los términos son estrictamente decrecientes, y $\frac{a_{n+2}}{a_{n+1}} < \frac{a_{n+1}}{a_n}$. Veamos esto:

$$\frac{a_{n+1}}{a_n} = \frac{\begin{pmatrix} R \\ n+1 \end{pmatrix} (t-1)^{R-n-1}}{\begin{pmatrix} R \\ n \end{pmatrix} (t-1)^{R-n}} \quad = \quad \frac{R-n}{(n+1)(t-1)}.$$

y reemplazando $n = \lfloor \frac{R}{t} + \theta \rfloor$, con $\theta > 0$

$$\begin{aligned} &= \frac{R - \lfloor \frac{R}{t} + \theta \rfloor}{(t-1)(\lfloor \frac{R}{t} + \theta \rfloor + 1)} \\ &\leq \frac{R - \frac{R}{t} - \theta + 1}{(t-1)(\frac{R}{t} + \theta)} \end{aligned}$$

$$\begin{aligned} &= \frac{(t-1)(\frac{R}{t} + \theta) - t\theta + 1}{(t-1)(\frac{R}{t} + \theta)} \\ &= 1 - \frac{(t\theta - 1)}{(t-1)(\frac{R}{t} + \theta)} \\ &< 1 - \frac{\theta}{\frac{R}{t} + \theta} \end{aligned}$$

El mayor de los cocientes de la forma $\frac{a_{n+1}}{a_n}$, cuando $n >= \lfloor \frac{R}{t} + \theta \rfloor$, y con $\varepsilon R/2 \leq \theta \leq R - \frac{R}{t}$ es: $\frac{a_{\lfloor \frac{R}{t} + \theta + 1}}{a_{\frac{R}{t} + \theta}}$ para $\theta = \frac{\varepsilon R}{2}$. Por lo tanto, usando la hipótesis $\varepsilon \leq \frac{1}{t}$ obtenemos:

$$1 - \frac{\theta}{\frac{R}{t} + \theta} = 1 - \frac{\varepsilon R/2}{\frac{R}{t} + \varepsilon R/2} = 1 - \frac{\varepsilon}{2/t + \varepsilon} \leq 1 - \frac{\varepsilon t}{3}$$

Ahora damos una cota superior para Z desplazando la suma a la izquierda $\lfloor \varepsilon R/2 \rfloor$ posiciones.

$$a_n = \frac{a_n}{a_{n-1}} \frac{a_{n-1}}{a_{n-2}} \cdots \frac{a_{n-\lfloor \varepsilon R/2 \rfloor + 1}}{a_{n-\lfloor \varepsilon R/2 \rfloor}} a_{n-\lfloor \varepsilon R/2 \rfloor} < (1 - \frac{\varepsilon t}{3})^{\lfloor \varepsilon R/2 \rfloor} a_{n-\lfloor \varepsilon R/2 \rfloor}$$

$$
\begin{aligned}
Z = \sum_{n=\lfloor \frac{R}{t}+\varepsilon R\rfloor}^{R} a_n &< \sum_{n=\lfloor \frac{R}{t}+\varepsilon R\rfloor}^{R} (1-\frac{\varepsilon t}{3})^{\lfloor \varepsilon R/2\rfloor} a_{n-\lfloor \varepsilon R/2\rfloor} \\
&= (1-\frac{\varepsilon t}{3})^{\lfloor \varepsilon R/2\rfloor} \sum_{n=\lfloor \frac{R}{t}+\varepsilon R\rfloor}^{R} a_{n-\lfloor \varepsilon R/2\rfloor} \\
&< (1-\frac{\varepsilon t}{3})^{\lfloor \varepsilon R/2\rfloor} t^R \ \ (porque \sum_{n=0}^{R} a_n = t^R)
\end{aligned}
$$

□

Ahora definiremos una serie de conjuntos que capturarán los números reales del intervalo (0,1) candidatos a ser absolutamente normales, y veremos algunas proposiciones que acotan las medidas de dichos conjuntos.

Recordemos la Def. 2, pero instanciada para dígitos en lugar de palabras. Sea α un real del intervalo $(0,1)$, d un dígito en base t. $S(\alpha,t,d,R)$ es el número de ocurrencias de d en los primeros R dígitos después del punto decimal, en la expresión de α en base t.

Diremos que s es candidata a ser absolutamente normal cuando la cantidad de apariciones de d en s (denotada con n) está "cerca"de la media.

Para probar que casi todos los reales en el intervalo $[0,1]$ son absolutamente normales, probaremos que hay "pocos" reales, en el sentido de la teoría de la medida, que son candidatos a no ser absolutamente normales. Toda vez que nos refiramos a la medida de un conjunto X, estaremos hablando de la medida de Lebesgue, y lo notaremos $\mu(X)$.

El resultado dado en el lema 8 muestra que podemos acotar superiormente la cantidad de palabras en base t de longitud R, que son candidatas a no ser absolutamente normales.

Definición 9. Sea $t \in \mathbb{N}$, $t > 1$, d un dígito en base t, $R \in \mathbb{N}, \varepsilon \in \mathbb{R}$.
Llamaremos $B(\varepsilon,d,t,R)$ al conjunto de reales $\alpha \in (0,1)$ tales que

$$|S(\alpha,t,d,R) - \frac{R}{t}| < \varepsilon R$$

Informalmente, los reales del conjunto $B(\varepsilon,\gamma,t,R)$ serán los candidatos a ser normales en base t. La siguiente Proposición acota inferiormente la medida de tales candidatos.

Proposición 10. *Sea $t \in \mathbb{N}$, $t > 1$, y sea d un dígito en base t. Para todo $R \in \mathbb{N}, \varepsilon \in \mathbb{R}$ tales que: $\frac{6}{R} \leq \varepsilon \leq \frac{1}{t}$:*

$$\mu\left(B(\varepsilon, d, t, R)\right) > 1 - 2(1 - \frac{\varepsilon t}{3})^{\lfloor \frac{R\varepsilon}{2} \rfloor}$$

Demostración. La medida de los $\alpha \in (0,1)$ tales que sus primeros R dígitos de su expresión fraccionaria en base t corresponden a una secuencia dada, es t^{-R}.

Luego,

$$\mu\left(\{\alpha \in (0,1) \colon |S(\alpha, t, d, R) - \frac{R}{t}| \geq \varepsilon R\}\right) = t^{-R} \sum_{|n - \frac{R}{t}| \geq \varepsilon R} N(t, d, n, R)$$

y por el Lema 8, tenemos:

$$\begin{aligned} \mu\left(B(\varepsilon, d, t, R)\right) &= \mu\left(\{\alpha \in (0,1) \colon |S(\alpha, t, d, R) - \frac{R}{t}| < \varepsilon R\}\right) \\ &> 1 - 2(1 - \frac{\varepsilon t}{3})^{\lfloor \frac{R\varepsilon}{2} \rfloor} \end{aligned}$$

□

Definimos $A(\varepsilon, T, R)$ como el conjunto de reales $\alpha \in (0,1)$ que son candidatos a ser normales en toda base t tal que $2 \leq t \leq T$:

Definición 11. Sea $T \in \mathbb{N}$, $T > 1$. Para todo $R \in \mathbb{N}, \varepsilon \in \mathbb{R}$:

$$A(\varepsilon, T, R) = \bigcap_{t=2}^{T} \bigcap_{d \in \{0, \ldots, t-1\}} B(\varepsilon, d, t, R)$$

En la siguiente Proposición acotamos inferiormente la medida de $A(\varepsilon, T, R)$:

Proposición 12. *Para todo $R \in \mathbb{N}, T \in \mathbb{N}, T > 1, \varepsilon \in \mathbb{R}$ tales que: $\frac{6}{R} \leq \varepsilon \leq \frac{1}{T}$:*

$$\mu\left(A(\varepsilon, T, R)\right) > 1 - \frac{T(T+1) - 2}{e^{\frac{R\varepsilon^2 - 2\varepsilon}{3}}}$$

Demostración. Tomando complemento, obtenemos las siguientes desigualdades:

$$\begin{aligned} \mu\left((0,1) \setminus A(\varepsilon, T, R)\right) &= \mu\left(\bigcup_{t=2}^{T} \bigcup_{d \in \{0, \ldots, t-1\}} (0,1) \setminus B(\varepsilon, d, t, R)\right) \\ &\leq \sum_{t=2}^{T} \sum_{d \in \{0, \ldots, t-1\}} \mu\left((0,1) \setminus B(\varepsilon, d, t, R)\right) \\ &< (T(T+1) - 2)(1 - \frac{2\varepsilon}{3})^{\lfloor \frac{R\varepsilon}{2} \rfloor} \\ &< (T(T+1) - 2)e^{-\frac{R\varepsilon^2 - 2\varepsilon}{3}} \end{aligned}$$

La anteúltima desigualdad surge de la Proposición 10 y de la observación de que el número de conjuntos $B(\varepsilon, d, t, R)$ que se intersecan es $\sum_{t=2}^{T} t = \frac{T(T+1)-2}{2}$. La última proviene del hecho que $(1+y/x)^x < e^y$, $\forall x > 0$. $\square$

Turing define los conjuntos A_k, especializando los conjuntos A para ciertos valores de ε, T y R. En esta sección, hacemos una variación a esas asignaciones, dado que estamos trabajando con otra cota inicial.

Definición 13. A_k es el conjunto $A(\varepsilon, T, R)$ especializando $\varepsilon = \frac{1}{\lfloor k^{1/4} \rfloor}$, $T = \lfloor k^{1/4} \rfloor$ y $R = k$. Es decir,

$$A_k = A(\frac{1}{\lfloor k^{1/4} \rfloor}, \lfloor k^{1/4} \rfloor, k)$$

Proposición 14. *$\exists k_0$ tal que, $\forall k \geq k_0$ tenemos*

$$\mu(A_k) \geq 1 - \frac{1}{k(k-1)}$$

Demostración. De la definición de A_k y por la Proposición 12, tenemos

$$\mu(A_k) = \mu\left(A(\frac{1}{\lfloor k^{1/4} \rfloor}, \lfloor k^{1/4} \rfloor, k)\right) > 1 - \frac{\lfloor k^{1/4} \rfloor(\lfloor k^{1/4} \rfloor + 1) - 2}{e^{\frac{k(\frac{1}{\lfloor k^{1/4} \rfloor})^2 - 2\frac{1}{\lfloor k^{1/4} \rfloor}}{3}}}$$

Se puede ver con facilidad que $\exists k_1 \in \mathbb{N}$ tal que, $\forall k \geq k_1$

$$\mu(A_k) > 1 - \frac{2\sqrt{k}}{e^{\frac{\sqrt{k}-2}{3}}}$$

Vemos ahora que $\exists k_0 \in \mathbb{N}$ que cumple que $\forall k \geq k_0$:

$$\frac{2\sqrt{k}}{e^{\frac{\sqrt{k}-2}{6}}} \leq \frac{1}{k(k-1)}$$

concluímos que: $\exists k_0$ tal que, $\forall k \geq k_0$, $\mu(A_k) \geq 1 - \frac{1}{k(k-1)}$ $\square$

Desde aquí, k_0 será el definido en la Proposición 14

Definición 15. Para cualquier natural $k \geq k_0$ y $n \geq 0$, $c(k, n) \subseteq (0, 1)$ se define de la siguiente manera:

$$c(k, n+1) = \begin{cases} (0,1) & \text{si } n = 0 \\ A_{k+n+1} \cap c(k,n) \cap (\beta_n, 1) & \text{si no} \end{cases}$$

donde β_n es elegido de forma tal que la medida de $c(k, n+1)$ es $1 - \frac{1}{k} + \frac{1}{k+n+1}$. Si $k < k_0$, definimos $c(k, n) = c(k_0, n)$.

De esta manera, $c(k, n)$ es una intersección finita de intervalos con extremos racionales, para cada k y n.

Observación 16. Para $n \geq 0$, siempre es posible encontrar β_n, pues

$$\mu\left(A_{k+n+1} \cap c(k,n)\right) \geq 1 - \frac{1}{k} + \frac{1}{k+n+1}.$$

para $k \geq k_0$. En efecto, si $n = 0$, $\mu\left(A_{k+n+1} \cap c(k,n)\right) = \mu\left(A_{k+n+1}\right) \geq 1 - \frac{1}{k(k+1)}$ la propiedad claramente vale. Para $n > 0$,

$$\begin{aligned}
\mu\left((0,1) \setminus (A_{k+n+1} \cap c(k,n))\right) &= \mu\left(((0,1) \setminus A_{k+n+1}) \cup ((0,1) \setminus c(k,n))\right) \\
&\leq \mu\left(((0,1) \setminus A_{k+n+1})\right) + \mu\left(((0,1) \setminus c(k,n))\right) \\
&\leq \frac{1}{(k+n+1)(k+n)} + \frac{1}{k} - \frac{1}{k+n} \\
&= \frac{1}{k} - \frac{1}{k+n+1}
\end{aligned}$$

de modo que

$$\mu\left(A_{k+n+1} \cap c(k,n)\right) \geq 1 - \frac{1}{k} + \frac{1}{k+n+1}.$$

Definición 17. Definimos $E_{c(k,n)}$ como el conjunto $c(k,n)$ habiendo removido los puntos extremos de cada intervalo.

Terminamos esta sección con la prueba del teorema principal, el Teorema 1 de Turing. Reescribimos el Teorema 1 con la función $E_{c(k,n)}$ introducida en la definición anterior. Aunque no es necesaria, la incluímos ya que Turing la utiliza en su demostración.

Demostración del Teorema 1 de Turing . Tenemos que ver que para $k \geq k_0$, el conjunto

$$E(k) = \bigcap_{n \geq 1} E_{c(k,n)}$$

tiene medida $1 - 1/k$ y contiene únicamente reales absolutamente normales.

La función $c(k,n)$ es recursiva, de modo que los conjuntos $E_{c(k,n)}$ también lo son.

Observemos que $\mu\left(E_{c(k,n)}\right) = 1 - \frac{1}{k} + \frac{1}{k+n+1}$, para $k \geq k_0$. $E_{c(k,n+1)} \subseteq E_{c(k,n)}$, de modo que

$$\begin{aligned}
\mu\left(E(k)\right) &= \lim_{n \to \infty} \mu\left(E_{c(k,n)}\right) \\
&= \lim_{n \to \infty} 1 - \frac{1}{k} + \frac{1}{k+n+1} \\
&= 1 - 1/k
\end{aligned}$$

Para ver que $E(k)$ sólo tiene reales absolutamente normales, vamos a ver que cualquier $\alpha \in E(k)(k > k_0)$ es simplemente normal para toda base t (ver nota al pie en la Definición 3).

Sea $\alpha \in E(k)$ y $d \in \{0, \ldots, t-1\}$.
Probaremos que:

$$\lim_{q\to\infty} \frac{S(\alpha, t, d, q)}{q} = \frac{1}{t}$$

$$\begin{aligned}
\alpha \in E(k) &\Rightarrow \alpha \in E_{c(k,n)} \forall n \in \mathbb{N}, n > 0 \\
&\Rightarrow \alpha \in c(k, n), \forall n(n > 0) \\
&\Rightarrow \alpha \in A_{k+n+1} \forall n > 0 \\
&\Rightarrow \alpha \in A(\frac{1}{\lfloor q^{1/4} \rfloor}, \lfloor q^{1/4} \rfloor, q), \forall q \geq k+2
\end{aligned}$$

Por la Definición 11, $\forall d\ (0 \leq d \leq t-1); \forall t(2 \leq t \leq \lfloor q^{1/4} \rfloor)$:
(Notar que $q > k_o \Rightarrow \frac{6}{q} \leq \varepsilon = \frac{1}{\lfloor q^{1/4} \rfloor}$

$$\alpha \in B(\frac{1}{\lfloor \sqrt{q} \rfloor}, d, t, q)$$

Por la Definición 9, $\forall d\ (0 \leq d \leq t-1); \forall t(2 \leq t \leq \lfloor q^{1/4} \rfloor)$:

$$|S(\alpha, t, d, q) - \frac{q}{t}| < \frac{q}{\lfloor q^{1/4} \rfloor}$$

$$\Rightarrow |\frac{S(\alpha, t, d, q)}{q} - \frac{1}{t}| < \frac{1}{\lfloor q^{1/4} \rfloor}$$

$$\Rightarrow \lim_{q\to\infty} |\frac{S(\alpha, t, d, q)}{q} - \frac{1}{t}| = 0$$

$\Rightarrow \alpha$ es simplemente normal en toda base $t : (2 \leq t \leq \lfloor q^{1/4} \rfloor), \forall q \geq k+2$
$\Rightarrow \alpha$ es absolutamente normal (definición 3).
Con esto queda demostrado el Teorema 1 de Turing.

3.1. Comparando las cotas

Nos interesa comparar las cotas del Misterioso Lema 6 y la obtenida en el Lema 8.

Sea $t \in \mathbb{N}$, $t > 1$. Sea d un dígito en base t. Para todo $R \in \mathbb{N}, \varepsilon \in \mathbb{R}$ tales que $\frac{6}{R} \leq \varepsilon \leq \frac{1}{t}$:

$$\sum_{|n - \frac{R}{t}| > \varepsilon R} N(t, d, n, R) \ < \ 2 \ (1 - \varepsilon t/3)^{\lfloor \varepsilon R/2 \rfloor} \, t^R.$$

Usando que $((1 + y/x)^x < e^y$, $\forall x > 0$, vemos que

$$\begin{aligned}
(1 - \varepsilon t/3)^{\lfloor \varepsilon R/2 \rfloor} &< e^{-\frac{\varepsilon t}{3} \lfloor \varepsilon R/2 \rfloor} \\
&< e^{-\frac{\varepsilon^2 Rt - 2\varepsilon t}{6}}
\end{aligned}$$

Si tomamos $k = \varepsilon R$, estamos sumando los mismos términos, y podemos ver que, si $\frac{kt}{R} < 0.3$:

$$\begin{aligned} 2\ (1-\varepsilon t/3)^{\lfloor \varepsilon R/2 \rfloor}\, t^R &< 2t^R e^{-\frac{k^2 t}{4R}} \\ \Leftrightarrow (1-\varepsilon t/3)^{\lfloor \varepsilon R/2 \rfloor} &< e^{-\frac{\varepsilon^2 R t}{4}} \end{aligned}$$

con lo que, la cota que obtuvimos es apenas mayor que la de Turing.

4. Versión fiel del Teorema 1 de Turing

En su manuscrito original Turing utiliza el siguiente resultado en la demostración del Teorema 1, y menciona que es posible probarlo, pero no da la demostración.

Misterioso Lema 18. *Sea $t \in \mathbb{N}$, $t > 1$. Sea γ una palabra en base t, con $\mid \gamma \mid = r$. Si $\frac{kt^r}{R} < 0.3$:*

$$\sum_{|n-Rt^{-r}|>k} N(t,\gamma,n,R) < 2t^R e^{-\frac{k^2 t^r}{4R}}$$

Dado que no pudimos reconstruir esta demostración, la asumiremos como hipótesis en esta sección.

Daremos ahora una versión extendida para palabras de cualquier longitud, de la Definición 9.

Definición 19. Sea $t \in \mathbb{N}$, $t > 1$, γ una palabra en base t con longitud r, $R \in \mathbb{N}, \Delta \in \mathbb{R}$.
Llamaremos $B(\Delta,\gamma,t,R)$ al conjunto de reales $\alpha \in (0,1)$ tales que

$$\mid S(\alpha,t,\gamma,R) - Rt^{-r} \mid < \frac{R}{\Delta t^r}$$

El primer paso en dirección a la demostración del Teorema 1 de Turing, es la siguiente proposición, análoga a la Proposición 10.

Proposición 20. *Sea $t \in \mathbb{N}$, $t > 1$, γ una palabra en base t con longitud r, $R \in \mathbb{N}, \Delta \in \mathbb{R}$. Si $\Delta^{-1} < 0.3$:*

$$m\ B(\Delta,\gamma,t,R) > 1 - 2e^{-\frac{Rt^{-r}}{4\Delta^2}}$$

Demostración. Siguiendo la misma idea que en la demostración de la Proposición 10, sabemos que:

$$\mu\left(B(\Delta,\gamma,t,R)\right) = 1 - t^{-R} \sum_{|n-Rt^{-r}|\geq \frac{R}{\Delta t^r}} N(t,\gamma,n,R)$$

Como $\Delta^{-1} < 0.3$, por la Proposición 18 tenemos

$$\begin{aligned} \mu\left(B(\Delta,\gamma,t,R)\right) &> 1 - t^{-R} 2t^R e^{-\frac{\frac{R^2}{\Delta^2 t^{2r}} t^r}{4R}} \\ &= 1 - 2e^{-\frac{Rt^{-r}}{4\Delta^2}}. \end{aligned}$$

□

La definición de A cambia levemente con respecto a la Definición 11

Definición 21. Sea $T \in \mathbb{N}$, $t > 1$, $L \in \mathbb{N}$, $R \in \mathbb{N}, \Delta \in \mathbb{R}$.

$$A(\Delta, T, L, R) = \bigcap_{t=2}^{T} \bigcap_{1 \leq |\gamma| \leq L} B(\Delta, \gamma, t, R)$$

Proposición 22. *La cantidad de conjuntos $B(\Delta, \gamma, L, R)$ que aparecen en la definición de $A(\Delta, T, L, R)$ es a lo sumo T^{L+1}.*

Demostración. No es difícil de ver que el número de conjuntos $B(\Delta, \gamma, L, R)$ que aparece en $\bigcap_{|\gamma| \leq L} B(\Delta, \gamma, t, R)$ es

$$\sum_{i=1}^{L} t^i = \frac{t^{L+1} - 1}{t - 1}.$$

Por lo tanto, el número de estos conjuntos en $A(\Delta, T, L, R) = \bigcap_{t=2}^{T} \bigcap_{|\gamma| \leq L} B(\Delta, \gamma, t, R)$ es

$$\sum_{t=2}^{T} \frac{t^{L+1} - 1}{t - 1} \leq T^{L+1}$$

(la última desigualdad sale sin problemas usando inducción en T). □

Probamos ahora un resultado similar a la Proposición 12:

Proposición 23. *Sea $T \in \mathbb{N}$, $t > 1$, $L \in \mathbb{N}$, $R \in \mathbb{N}, \Delta \in \mathbb{R}$.*
Si $\Delta^{-1} < 0.3$:

$$\mu\left(A(\Delta, T, L, R)\right) > 1 - 2T^{L+1} e^{-\frac{RT^{-L}}{4\Delta^2}}$$

Demostración. Tal como hicimos en la demostación de la Proposición 12, tomamos complemento, y obtenemos las siguientes desigualdades:

$$\begin{aligned}
\mu\left((0,1) \setminus A(\Delta, T, L, R)\right) &= \mu\left(\bigcup_{t=2}^{T} \bigcup_{1 \leq |\gamma| \leq L} (0,1) \setminus B(\Delta, \gamma, t, R)\right) \\
&\leq \sum_{t=2}^{T} \sum_{1 \leq |\gamma| \leq L} \mu\left((0,1) \setminus B(\Delta, \gamma, t, R)\right) \\
&< 2T^{L+1} e^{-\frac{RT^{-L}}{4\Delta^2}}.
\end{aligned}$$

La última desigualdad sale de las proposiciones 20 y 22. □

Tal como hicimos en la Definición 13, asignamos valores para las variables Δ, T, L y R en $A(\Delta, T, L, R)$. Turing utiliza otras asignaciones, ($\Delta = \lfloor k^{1/4} \rfloor, T = \lfloor e^{\sqrt{lnk}} \rfloor, L = \lfloor \sqrt{lnk} - 1 \rfloor, R = k$). Esto se debe a que existen pequeñas diferencias en la cota obtenida para $\mu\left(A(\Delta, T, L, R)\right)$ con respecto a la de Turing. Sin embargo, las asignaciones respetan que todos los parámetros crecen con k y recorren todos los valores posibles:

Definición 24. A_k es el conjunto $A(\Delta, T, L, R)$ especializando $\Delta = \lfloor k^{1/8} \rfloor$, $T = \lfloor (\ln k)^{1/4} \rfloor$, $L = \lfloor (\ln k)^{1/4} - 1 \rfloor$ y $R = k$. Es decir,

$$A_k = A(\lfloor k^{\frac{1}{8}} \rfloor, \lfloor (\ln k)^{1/4} \rfloor, \lfloor (\ln k)^{1/4} - 1 \rfloor, k)$$

Con estas especializaciones de Δ, T, L y R, podemos probar un resultado similar a la Proposición 14.

Proposición 25. *Existe un k_0 tal que $\forall k \geq k_0$*

$$\mu(A_k) \geq 1 - \frac{1}{k(k-1)}$$

Demostración. De la Definición 24 y por la Proposición 23, tenemos que para k suficientemente grande

$$\begin{aligned}
\mu(A_k) &= \mu\left(A(\lfloor k^{\frac{1}{8}} \rfloor, \lfloor (\ln k)^{1/4} \rfloor, \lfloor (\ln k)^{1/4} - 1 \rfloor, k)\right) \\
&> 1 - \frac{2\lfloor (\ln k)^{1/4} \rfloor^{\lfloor (\ln k)^{1/4} - 1 \rfloor + 1}}{e^{\frac{k}{4\lfloor k^{\frac{1}{8}} \rfloor^2 \lfloor (\ln k)^{1/4} \rfloor^{\lfloor (\ln k)^{1/4} - 1 \rfloor}}}} \\
&> 1 - \frac{2\sqrt{k}}{e^{\frac{k^{1/4}}{4}}}
\end{aligned}$$

También se puede ver que para k suficientemente grande,

$$1 - \frac{2\sqrt{k}}{e^{\frac{k^{1/4}}{4}}} > 1 - \frac{1}{k(k-1)}.$$

□

Las definiciones 15 y 17 y la observación 16 se siguen cumpliendo.

Demostración del teorema 1 de Turing. Tenemos que ver que para $k \geq k_0$, el conjunto

$$E(k) = \bigcap_{n \geq 1} E_{c(k,n)}$$

tiene medida $1 - 1/k$ y contiene únicamente reales normales (de acuerdo a la Definición 4 que usa Turing).

Para ver que este conjunto tiene medida $1 - 1/k$, usamos los mismos argumentos que vimos en la sección anterior.

Para ver que $E(k)$ sólo tiene reales normales, vamos a ver que cualquier $\alpha \in E(k)(k > k_0)$ es normal para toda base t.

Sea $\alpha \in E(k)$ y γ una palabra en base t, con $\mid \gamma \mid= r$. Probaremos que

$$\lim_{q\to\infty} \frac{S(\alpha,t,\gamma,q)}{q} = \frac{1}{t^r}$$

$$\begin{aligned}
\alpha \in E(k) \quad &\Rightarrow \quad \alpha \in E_{c(k,n)} \forall n \in \mathbb{N}, n > 0 \\
&\Rightarrow \quad \alpha \in c(k,n), \forall n(n > 0) \\
&\Rightarrow \quad \alpha \in A_{k+n+1} \forall n > 0 \\
&\Rightarrow \quad \alpha \in A(\lfloor q^{\frac{1}{8}} \rfloor, \lfloor (\ln q)^{1/4} \rfloor, \lfloor (\ln q)^{1/4} - 1 \rfloor, q), \forall q \geq k+2
\end{aligned}$$

Por la Definición 21,
$\forall t(2 \leq t \leq \lfloor (\ln q)^{1/4} \rfloor); \forall \gamma$ palabra en base t con $\mid \gamma \mid= r$ y $r \leq \lfloor (\ln q)^{1/4} - 1 \rfloor$:
(Notar que $q > k_o \Rightarrow \frac{1}{\Delta} \leq 0.3$)

$$\alpha \in B(\frac{1}{\lfloor q^{1/8} \rfloor}, \gamma, t, q)$$

Por la Definición 19,
$\forall t(2 \leq t \leq \lfloor (\ln q)^{1/4} \rfloor); \forall \gamma$ palabra en base t con $\mid \gamma \mid= r$:

$$|S(\alpha,t,\gamma,q) - \frac{q}{t^r}| < \frac{q}{\lfloor q^{1/8} \rfloor t^r} < \frac{q}{\lfloor q^{1/8} \rfloor}$$

$$\Rightarrow |\frac{S(\alpha,t,\gamma,q)}{q} - \frac{1}{t^r}| < \frac{1}{\lfloor q^{1/8} \rfloor}$$

$$\Rightarrow \lim_{q\to\infty} |\frac{S(\alpha,t,\gamma,q)}{q} - \frac{1}{t}| = 0$$

$\Rightarrow \alpha$ es normal (Definición 4) en toda base $t : (2 \leq t \leq \lfloor (\ln q)^{1/4} \rfloor), \forall q \geq k+2$.

Queda demostrado el Teorema 1 de Turing.

5. Una cota alternativa

En esta sección, presentaremos una cota alternativa a la dada en el Lema 8. Este camino, fue propuesto por la Dra. Mariela Sued, a quien le agradecemos enormemente su aporte.

Si bien la cota que daremos aquí, no es más fina que la que utilizamos en la Sección 3, la incluiremos para dar otro enfoque al mismo problema y que también nos permite hacer la reconstrucción del Teorema 1 deTuring, en una manera similar a la realizada en la sección 3.

La notación utilizada en esta seccíon, es la siguiente: $P(x)$ es la función de Probabilidad, $E(x)$ es la esperanza matemática, y B(n,p) representa la función de distribución binomial, donde n representa el número de pruebas y p la probabilidad de éxito en cada prueba.

Para desarrollar esta cota, utilizaremos el siguiente resultado, cuya demostración se encuentra en el libro de D.Pollard [8] (Hoeffding's Inequality y Corollary 3, pag 191/192):

Proposición 26 (Desigualdad de Hoeffding). *Sean $Y_1, Y_2, ..., Y_R$ variables aleatorias independientes, con media 0 y rangos acotados: $a_i \leq Y_i \leq b_i$. Para cada $n > 0$*

$$P(Y_1 + ... + Y_R \geq n) \leq e^{-(2n^2/\sum_{i=1}^{R}(b_i - a_i)^2)}$$

Corolario 27. *Bajo las mismas condiciones que la Proposición anterior, se cumple que:*

$$P(\mid Y_1 + ... + Y_R \mid \geq n) \leq 2e^{-(2n^2/\sum_{i=1}^{R}(b_i - a_i)^2)}$$

La siguiente proposición, demuestra la cota alternativa, que presentamos en esta sección:

Proposición 28. *Sea $t \in \mathbb{N}$, $t > 1$. Sea d un dígito en base t.*

$$\sum_{|n - \frac{R}{t}| > k} N(t, d, n, R) < 2t^R e^{-\frac{2k^2}{R}}$$

Demostración. Se puede ver que: $\frac{N(t,d,n,R)}{t^R}$ tiene distribución Binomial de parámetros: $B(R, 1/t)$. En este modelo, cada prueba representa la ocurrencia o no de d, en cada uno de los R dígitos de cada palabra. De esta manera, las pruebas son independientes, (el hecho que aparezca d en una posición dada, no afecta la probabilidad de que aparezca en las siguientes)y la probabilidad de éxito de cada prueba es $1/t$.

Sean X_i variables aleatorias independientes, con distribución Binomial de parámetros: $B(1, 1/t)$. De esta manera, vemos que: $\sum_{i=1}^{R} X_i = n$ y $E(X_i) = 1/t$
Cada X_i representa a cada una de las R pruebas involucradas en $\frac{N(t,d,n,R)}{t^R}$
Ahora definimos: Y_i variables aleatorias independientes, de la siguiente manera: $Y_i = X_i - 1/t$
Observemos que:
$a_i = -1/t \leq Y_i \leq b_i = \frac{t-1}{t}$ y $E(Y_i) = (-1/t)\frac{t-1}{t} + 1/t\frac{t-1}{t} = 0$

Vemos que: $\sum_{i=1}^{R} Y_i = \sum_{i=1}^{R}(X_i - 1/t) = n - \frac{R}{t}$ y también que: $E(\sum_{i=1}^{R} Y_i) = \sum_{i=1}^{R} E(Y_i) = 0$ y $\sum_{i=1}^{R}(b_i - a_i)^2 = R$

Usando el Corolario 27 (que se deduce la desigualdad de Hoeffding, Prop.26), obtenemos:
$P(\mid n - \frac{R}{t} \mid > k) \leq 2e^{\frac{-2k^2}{R}}$
con lo que: $\sum_{|n-\frac{R}{t}|>k} N(t, d, n, R) < 2t^R e^{-\frac{2k^2}{R}}$

$\square$

A continuación, veremos que esta cota es mayor que la que la obtenida en el lema 8, (y por lo tanto, mayor que la utilizada por Turing, como vimos en la sección 3). Recordemos la cota del Lema 8:

Sea $t \in \mathbb{N}$, $t > 1$. Sea d un dígito en base t. Para todo $R \in \mathbb{N}, \varepsilon \in \mathbb{R}$ tales que $\frac{6}{R} \leq \varepsilon \leq \frac{1}{t}$:

$$\sum_{|n-\frac{R}{t}|>\varepsilon R} N(t, d, n, R) \ < \ 2 \ (1 - \varepsilon t/3)^{\lfloor \varepsilon R/2 \rfloor} \, t^R.$$

Si tomamos $k = \varepsilon R$, estamos sumando los mismos términos, y podemos ver que:

$$2 \ (1 - \varepsilon t/3)^{\lfloor \varepsilon R/2 \rfloor} \, t^R < 2t^R e^{-\frac{2k^2}{R}} \Leftrightarrow (1 - \varepsilon t/3)^{\lfloor \varepsilon R/2 \rfloor} < e^{-\frac{2k^2}{R}}$$

Luego, usando que $((1 + y/x)^x < e^y$, $\forall x > 0$, vemos que:

$$(1 - \varepsilon t/3)^{\lfloor \varepsilon R/2 \rfloor} < e^{-\frac{\varepsilon t}{3}\lfloor \varepsilon R/2 \rfloor} < e^{-\frac{\varepsilon^2 Rt - 2\varepsilon t}{6}}$$

Como $k = \varepsilon R$, vemos que:

$$e^{-\frac{\varepsilon^2 Rt-2\varepsilon t}{6}} < e^{-\frac{2k^2}{R}} \Leftrightarrow e^{-\frac{\varepsilon^2 Rt-2\varepsilon t}{6}} < e^{-2\varepsilon^2 R} \Leftrightarrow \frac{\varepsilon^2 Rt - 2\varepsilon t}{6} > 2\varepsilon^2 R \Leftrightarrow \varepsilon R > \frac{2t}{t - 12}$$

Y como la cota del lema 8, vale para $\varepsilon \geq 6/R$, esta cota es mejor para todo t tal que:

$$6 > \frac{2t}{t - 12}, \text{ o sea,} \forall t > 18$$

Por lo tanto, para la mayor parte de los casos, la cota del Lema 8 se comporta mejor que la dada aquí.

6. Conclusiones

Si bien el resultado de este teorema no es novedoso en si mismo, si es novedosa la versión constructiva de esta demostración. A pesar de que no hemos logrado reconstruir fielmente el teorema, si podemos concluír que la estrategia utilizada es correcta.

Por otro lado, vimos que los aportes del editor son insuficientes para la comprensión del trabajo de Turing ya que no demuestra ningún resultado, e inclusive, algunas notas son incorrectas.

Con este trabajo, creemos haber aclarado un poco las estrategias utilizadas por Turing en la demostración del Teorema 1. Queda por ver el Teorema 2, e incluso se podría intentar demostrar el misterioso lema en su versión basada en palabras.

Referencias

[1] V. Becher, S. Figueira. An example of a computable absolutely normal number, *Theoretical Computer Science*, Vol.270, pp. 947-958, 2002.

[2] E. Borel. Les probabilités dénombrables et leurs applications arithmetiques. *Rend. Circ. Mat. Palermo*, 27:247–271, 1909.

[3] David H. Bailey and Richard E. Crandall. On the Random Character of Fundamental Constant Expansions, *Experimental Mathematics*, vol. 10, no. 2 , pp. 175-190, 2001.

[4] David H. Bailey and Richard E. Crandall. Random Generators and Normal Numbers, *Experimental Mathematics*, vol. 11, no. 4 , pp 527-546, 2004.

[5] J. Borwein, D. Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century.* Natick, MA: A. K. Peters, p. 143, 2003.

[6] G. Chaitin. A Theory of Program Size Formally Identical to Information Theory, *Journal of the ACM* (JACM), v.22 n.3, p.329-340, July 1975.

[7] G.Harman. *Metric Number Theory.* London Mathematical Society Monographs. Oxford Universaity Press, 1998.

[8] D. Pollard, *Convergence of Stochastic Processes.* Springer Verlag New York, 1984.

[9] M. W. Sierpinski. Démonstration élèmentaire du théorème de M. Borel sur les nombres absolument normaux et détermination effective d'un tel nombre. *Bull. Soc. Math. France*, 45:127–132, 1917.

[10] M. W. Sierpinski. *Elementary Theory of Numbers*, Warszawa, 1964.

[11] A. Turing. A Note on Normal Numbers. *Collected Works of Alan M. Turing, Pure Mathematics*, edited by J. L. Britton, pp. 117-119. North Holland, 1992.

7. Apéndice 1
Manuscritos Originales de A.M.Turing

Page1

Page2

Page3

Page4

Page5

Page6

Page7

Page8

Page9

Page10

Page11

Page12

Page13

Page14

Page15

Page16

Page17

Page18

8. Apéndice 2
Versión transcripta por J.L.Britton