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.

Al codirector de mi tesis, 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

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* $k \geq k_0$*:* $c(k, n+1) \subsetneq c(k, n)$ *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 que utilizamos en nuestra reconstrucción del Teorema 1 de Turing.

En los apéndices 1 y 2, incluimos 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 referirnos 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 7).

Definición 4. Sea α un número real en $(0,1)$.
α es *normal* si para toda base t, y 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 esta 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,
$N(t,d,n,R) = \left(\begin{array}{c} R \\ n \end{array} \right) (t-1)^{R-n}$

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.

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

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

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

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

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

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

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

$$\begin{aligned} &\text{para } \theta \leq \tfrac{R}{t} \\ &\qquad < 1 - \frac{t(\theta-1)}{R} \\ &\text{para } \theta \geq \varepsilon R/2 \\ &\qquad \leq 1 - \frac{\varepsilon t}{2} + \frac{t}{R} \end{aligned}$$

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 $0 \leq n \leq \frac{R}{t} - \varepsilon R$,

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

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

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. Sabemos que $\forall n > \lfloor \frac{R}{t} \rfloor + 1$, $a_n > a_{n+1}$, o sea que los términos son estrictamente decrecientes, y $\frac{a_{n+2}}{a_{n+1}} < \frac{a_{n+1}}{a_n}$. Demos una cota para estos cocientes:

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

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

$$= \frac{R - \lfloor \frac{R}{t} + \theta \rfloor}{(t-1)(\lfloor \frac{R}{t} + \theta \rfloor + 1)}$$

$$\leq \frac{R - \frac{R}{t} - \theta + 1}{(t-1)(\frac{R}{t} + \theta)}$$

$$= \frac{(t-1)(\frac{R}{t} + \theta) - t\theta + 1}{(t-1)(\frac{R}{t} + \theta)}$$

$$= 1 - \frac{(t\theta - 1)}{(t-1)(\frac{R}{t} + \theta)}$$

$$< 1 - \frac{\theta}{\frac{R}{t} + \theta}$$

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: $\frac{a_{\lfloor \frac{R}{t} + \theta + 1 \rfloor}}{a_{\lfloor \frac{R}{t} + \theta \rfloor}}$ 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: $\frac{6}{R} \leq \varepsilon \leq \frac{1}{T}$:*

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

Demostración. Tomando complemento, obtenemos las siguientes desigualdades:

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

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 $\forall x > 0$ y $\forall y: \ (1+y/x)^x < e^y$. □

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

$$\mu(A_k) = \mu\left(A(\frac{1}{\lfloor k^{1/4} \rfloor}, \lfloor k^{1/4} \rfloor, k)\right) > 1 - \frac{\lfloor k^{1/4} \rfloor(\lfloor k^{1/4} \rfloor + 1) - 2}{e^{k(\frac{1}{\lfloor k^{1/4} \rfloor})^2 - 2\frac{1}{\lfloor k^{1/4} \rfloor}}_3}$$

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

$$\mu(A_k) > 1 - \frac{2\sqrt{k}}{e^{\sqrt{k}-2}_3}$$

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

$$\frac{2\sqrt{k}}{e^{\sqrt{k}-2}_6} \leq \frac{1}{k(k-1)}$$

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

Desde aquí, k_0 será el determinado 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 un número racional, 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 $n > 0$,

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

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

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

La cota de Turing, es apenas menor que la cota obtenida del Lema 8, ya que:

$$2t^R e^{-\frac{k^2 t}{4R}} < 2t^R e^{-\frac{\varepsilon^2 Rt - 2\varepsilon t}{6}}$$

si tomamos $k = \varepsilon R$, porque

$$2t^R e^{-\frac{\varepsilon^2 Rt}{4}} < 2t^R e^{-\frac{\varepsilon^2 Rt - 2\varepsilon t}{6}}$$

Turing da su cota para $\frac{kt}{R} < 0.3$. Tomando $k = \varepsilon R$, dice $\varepsilon < \frac{0.3}{t}$, mientras que nuestra cota, pide $\frac{6}{R} \leq \varepsilon < \frac{1}{t}$.

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, ($\Delta = \lfloor k^{1/4} \rfloor, T = \lfloor e^{\sqrt{ln\, k}} \rfloor, L = \lfloor \sqrt{ln\, k} - 1 \rfloor, R = k$) que comprobamos que no son apropiadas. Sin embargo, las asignaciones que utilizamos respetan que todos los parámetros crecen con k y recorren todos los valores posibles, al igual que las utilizadas por Turing:

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

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

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

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

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

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 $\mid \gamma \mid = r$:

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

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

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

$\Rightarrow \alpha$ 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 k real, $k > 0$*

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

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

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

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

□

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

El interés del Teorema 1 es el hecho de que da una construcción recursiva para demostrar que el conjunto de números normales tiene medida de Lebesgue igual a 1.

A pesar de que no hemos conseguido reconstruir la demostración del teorema con la estrategia original de Turing, nuestra estrategia también da una construcción recursiva.

Cabe señalar que las notas del editor ([11], pp. 263-365) sobre el manuscrito original son insuficientes para la comprensión del trabajo, inclusive, hay algunas notas incorrectas.

Con esta tesis, hemos conseguido identificar las herramientas utilizadas por Turing en la demostración del Teorema 1. Creemos que es posible dar una demostración del Misterioso Lema en su versión basada en palabras, cambiando levemente la cota dada por Turing. Y por supuesto, queda por estudiar el Teorema 2, el algoritmo para generar números normales, que utiliza la construcción del Teorema 1.

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

A Note on Normal Numbers

A. M. Turing

Although it is known that all numbers are normal 1) no example of a normal number has ever been given. I propose to show how normal numbers can be constructed and ~~[illegible]~~ by ~~[illegible]~~ to give a rule by which for each $\varepsilon > 0$ normal numbers are computed between 0 and 1, the measure of the set of these normal numbers exceeding $1-\varepsilon$.

It may be noticed that ~~[illegible]~~ ~~[illegible]~~ examples of numbers normal in the ~~scale of 10~~ have clearly been given by Champernowne 2). It may also be noticed that ~~[illegible]~~ the ~~examples of~~ computable normal numbers ~~given~~ have not actually been written down; ~~a further [illegible]~~. ~~the difficulty of~~ This rule cannot therefore be considered as providing concrete examples of normal numbers, but rather as a counter ~~argument~~ to the opinion that the existence proof for normal numbers provides no examples of them.

The argument ~~in~~ the rule ~~which~~ follows ~~quite closely the~~ the existence proof fairly closely.

of steps. When this figure has been calculated and written down as the $R(N)$ th figure of β', the N th section is finished. Hence $\mathcal{H}$ is circle-free.

Now let K be the D.N of $\mathcal{H}$. What does $\mathcal{H}$ do in the K th section of its motion? It must test whether K is satisfactory giving a verdict 's' or 'u'. Since K is the D.N of $\mathcal{H}$ and since $\mathcal{H}$ is circle-free, the verdict cannot be 'u'. On the other hand the verdict cannot be 's'. For if it were, then in the K th section of its motion $\mathcal{H}$ would be bound to compute the first $R(K-1)+1 : R(K)$ figures of the sequence computed by the machine with K as its D.N and to write down the $R(K)$ th as a figure of the sequence computed by $\mathcal{H}$. The computation of the first $R(K)-1$ figures would be carried out all right, but the instructions for calculating the $R(K)$ th would amount to "calculate the first $R(K)$ figures computed by $\mathcal{H}$ and write down the $R(K)$th". This $R(K)$ th figure would never be found. i.e., $\mathcal{H}$ is circular, contrary both to what we have found in the last paragraph and to the verdict 's'. Thus both verdicts are impossible and we conclude that there can be no machine $\mathcal{D}$.

We can show further that <u>there can be no machine $\mathcal{E}$ which, when supplied with the S.D of an arbitrary machine $\mathcal{M}$, will determine whether $\mathcal{M}$ ever prints a given symbol (0 say)</u>.

We will first show that if there is a machine $\mathcal{E}$ then there is a general process for determining whether a given machine $\mathcal{M}$ prints 0 infinitely often. Let $\mathcal{M}_1$ be a machine which prints the same sequence as $\mathcal{M}$, except that in the position where the first 0 printed by stands, $\mathcal{M}_1$, prints 0. $\mathcal{M}_2$ is to have the first two symbols 0 replaced by $\overline{0}$, and so on. Thus if $\mathcal{M}$ were to print

A B A 0 1 A A B 0 0 1 0 A B . . .

If γ is a sequence of digits in the scale of t let

If γ is a sequence of digits in the scale of t we

<u>Definition</u> $l(\gamma)$ is the length of γ.

is the length of γ.

If α is a real number between 0 and 1 and γ is a sequence of digits in the scale of t we put

$S(\alpha, t, \gamma, R) =$ number of occurrences of γ in the first R figures of the decimal sequence of α in the scale of t.

α is said to be normal in the scale of t with respect to γ if

$$R^{-1} S(\alpha, t, \gamma, R) \to t^{-l(\gamma)} \quad \text{as } R \to \infty \qquad (1)$$

It is said to be normal in the scale of t if (1) holds for all γ, and to be normal if it holds for all γ, t.

We shall also put

$N(n, t, \gamma, R) =$ number of sequences of R digits in scale of t in which γ occurs exactly n times.

Instead of saying α is normal with respect to γ we may say that γ has normal frequency in α.

figure is $\beta_R(N)$.

... we can invent a machine which when supplied with the D.N of any computing machine ... will test this D.N. and if ... is circular will mark the D.N. with the symbol ... circle free ... will mark it ... and ... we could construct a machine $\mathcal{H}$ to compute the sequence β. The machine ... may require a tape. ... beyond all symbols on F-squares, and that when it has reached its verdict all the rough work done by ...

... has its motion divided into sections. In the first $N-1$ sections among other things the integers $1, 2, \ldots, N-1$ have been written down and tested by the machine ... A certain number, say $R(N-1)$ of them have been found to be the D.N's of circle free machines. In the Nth section the machine ... tests the number N. If N is satisfactory i.e., if it is the D.N. of a circle free machine, then $R(N) = 1 + R(N-1)$ and the first $R(N)$ figures of the sequence of which a D.N. is N are calculated. The $R(N)$th figure of this sequence is written down as one of the figures of the sequence ... computed by $\mathcal{H}$. If N is not satisfactory, then $R(N) = R(N-1)$ and the machine goes on to the $(N+1)$th section of its motion.

From the construction of $\mathcal{H}$ we can see that $\mathcal{H}$ is circle free. Each section of the motion of $\mathcal{H}$ comes to an end after a finite number of steps. For by our assumption ... the decision as to whether N is satisfactory is reached in a finite number of steps. If N is not satisfactory then the Nth section is finished. If N is satisfactory this means that the machine $M(N)$ whose D.N. is N is circle free, and therefore its $R(N)$th figure can be calculated in a finite number

figure is $\varphi_n(n)$.

Let us suppose that there is such a process; that is to say that we can invent a machine $\mathscr{D}$ which, when supplied with the S.D of any computing machine $\mathscr{M}$ will test this S.D. and if $\mathscr{M}$ is circular will mark the S.D. with the symbol 'u' and if it is ~~vicious~~ circle free will mark it with 's'. By combining the machines $\mathscr{D}$ and $\mathscr{U}$ we could construct a machine $\mathscr{H}$ to compute the sequence β'. The machine $\mathscr{D}$ may require a tape. We may suppose that it uses the E-squares beyond all symbols on F-squares, and that when it has reached its verdict all the rough work done by $\mathscr{D}$ is erased.

The machine $\mathscr{H}$ has its motion divided into sections. In the first $N-1$ sections amongst other things the integers $1, 2, \ldots, N-1$ have been written down and tested by the machine $\mathscr{D}$. A certain number, say $R(N-1)$ of them have been found to be the D.N.'s of circle-free machines. In the Nth section the machine $\mathscr{D}$ tests the number N. If N is satisfactory i.e., if it is the D.N. of a circle free machine, then $R(N) = 1 + R(N-1)$ and the first $R(N)$ figures of the sequence of which a D.N. is N are calculated. The $R(N)$th figure of this sequence is written down as one of the figures of the sequence β' computed by $\mathscr{H}$. If N is not satisfactory, then $R(N) = R(N-1)$ and the machine goes on to the $(N+1)$th section of its motion.

From the construction of $\mathscr{H}$ we can see that $\mathscr{H}$ is circle free. Each section of the motion of $\mathscr{H}$ comes to an end after a finite number of steps. For by our assumption about $\mathscr{D}$ the decision as to whether N is satisfactory is reached in a finite number of steps. If N is not satisfactory then the Nth section is finished. If N is satisfactory this means that the machine $\mathscr{M}(N)$ whose D.N. is N is circle free, and therefore its $R(N)$ th figure can be calculated in a finite number

Summary!

If a number is normal in the scale of b, w.r.t. ~~every admissible~~ sequence which are not of form $\beta_1 \beta_2 \beta_3$ then it is normal in the scale of b^t.

If not normal let us consider one of the shortest sequences β_1 in a case $\beta_1 \beta_2 \beta_3$ of failure. Let β_2 be [illegible] possible consistent with β_1 having the same length and ~~of~~ ~~[illegible]~~ let the number of repetitions of β_1 at the end of β_2 also have the minimum value. The resulting sequence is (k times (say))

$$\beta_1 \beta_3 \beta_1 \beta_1 \dots \beta_1$$

Now $\beta_1 \beta_3$ ~~is not normal~~ has normal frequency ~~[illegible]~~ ~~[illegible]~~ being too short. The others in, and so have every sequence $\beta_1 \beta_3 \beta_4$ where β_4 is different to $\beta_1 \dots \beta_1$ (k times) (say) but of the same length. In $\beta_1 \beta_3 \beta_4$ [illegible] has too repetitions of β_1, [illegible] be abnormal [illegible] normal freq. Then by difference $\beta_1 \beta_3 \beta_1 \dots \beta_1$ has normal freq, contrary to hyp.

(vi) If α and β are computable and $\alpha < \beta$ and $\varphi(\alpha) < 0 < \varphi(\beta)$ where $\varphi(\alpha)$ is a computable increasing continuous function, then there is a unique computable number γ, satisfying $\alpha < \gamma < \beta$ and $\varphi(\gamma) = 0$.

Computable Convergence

We will say that a sequence β_n of computable numbers converges computably if there is a computable integral valued function $N(\varepsilon)$ of the computable variable ε, such that we can show that if $\varepsilon > 0$ and $n > N(\varepsilon)$ and $m > N(\varepsilon)$, than $|\beta_n - \beta_m| < \varepsilon$.

We can then shew

(vii) A power series whose coefficients form a computable sequence of computable numbers is computably convergent in the of its interval of convergence.

(viii) The limit of a computably convergent sequence is computable. And with the obvious definition of "uniformly computably convergent"

(ix) The limit of a uniformly computably convergent computable sequence of computable functions is a computable function. Whence

(x) The sum of a power series whose coefficients form a computable sequence is a computable function in the interior of its interval of convergence.

From (viii) and $\overline{\pi} = 4\left(1 - \frac{1}{3} + \frac{1}{5} \cdots\right)$ we deduce that $\overline{\pi}$ is computable.

From $e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} \cdots$ we deduce that e is computable.

Lemma 6.

If γ is a sequence each of the $\beta_1 \beta_2 \beta_3$ each of length l then the number $N(k, \gamma, n, R)$ of ~~[illegible]~~ which of length R in which γ occurs exactly n times satisfies

$$N(k, \gamma, n, R) \leq \frac{(R/k^l)^n}{n!} k^R$$

$N(k, \gamma, n, R)$ = No. of ways of choosing positions of sequences γ × No. of ways of filling in remaining terms

$$\leq \frac{R(R-l)(R-2l)\ldots(R-nl+l)}{n!} k^{R-nl}$$

$$\leq \frac{(R/k^l)^n}{n!} k^R$$

~~$$\sum_{|n - R/k^l| > K} N(k, \gamma, n, R) \leq k^R \left[\sum_{n > R/k^l + K} \frac{(R/k^l)^n}{n!} \right.$$~~

$$\sum_{n < R/k^l - K} \frac{(R/k^l)^n}{n!}$$

~~$$\leq k^R \left[\frac{(R/k^l)^{R/k^l}}{\Gamma(R/k^l + 1)} \cdot \frac{\Gamma(R/k^l + 1)}{\Gamma(n+1)} \right] \bigg/ \left(1 - \frac{R/k^l}{R/k^l + K}\right)$$~~

(iv) If $\varphi(n)$ is a computable function whose value is always 0 or 1, then the sequence whose n^{th} figure is $\varphi(n)$ is computable.

Dedekind's theorem does not hold in the ordinary form if we replace "real" throughout by "computable". But it holds in the following form

(v) If $G(\alpha)$ is a propositional function of the computable numbers and

$$(a)\quad (\exists \alpha)(\exists \beta)\{G(\alpha) \neq (-G(\beta))\}$$

$$(b)\quad G(\alpha) \neq (-G(\beta)) \rightarrow (\alpha < \beta)$$

and there is a general process for determining the truth value of $G(\alpha)$ then there is a computable number ξ such that

$$G(\alpha) \rightarrow \alpha \leq \xi$$

$$-G(\alpha) \rightarrow \alpha \geq \xi$$

In other words the theorem holds for any section of the computables such that there is a general process for determining to which class a given number belongs.

Owing to this restriction of Dedekind's Theorem we cannot say that a computable bounded increasing sequence of computable numbers has a computable limit. This may possibly be understood by considering a sequence such as,

$$-1, -\tfrac{1}{2}, -\tfrac{1}{4}, -\tfrac{1}{8}, -\tfrac{1}{16}, \tfrac{1}{2}, \ldots$$

On the other hand (v) enables us to prove

5

Let γ be distinct fixed but quadratically ... K of

$N(k, \gamma, n, R)$ = No of R-figure ... in the scale of 1, 2 ...

in which ... occurs exactly n times.

We have

$N(k, \gamma, n, R)$ = No of ways of placing the ~~figures~~ figures ...
× No of ways of filling in the gaps.

$$\leq \frac{R(R-k)(R-2k)\cdots(R-\overline{n-1}k)}{\lfloor n} \, t^{R-nk}$$

$$\leq \frac{R(R-1)\cdots(R-k)}{\lfloor n} \, t^{R-nk}$$

$$= \binom{R}{n} t^{R-nk}$$

= coeff of z^n in $(1+z/t^k)^R$

$t^R (1+z/t^k)^R$

$(t+zt^{1-k})^R$

$(z+t)^R$

and

$$\mathfrak{A}_\eta \& F^{(M)} \to \Big[F\big(u^{(u-1)}, u^{(u)}\big) \& H\big(u^{(u-1)}, u^{(\eta(u-1))}\big) \& K\big(u^{(u)}, u^{(\eta(u-1))}, u^{(\eta(u))}\big) \to H\big(u^{(u)}, u^{(\eta(u))}\big)$$

Whence

$$\mathfrak{A}_\eta \& F^{(M)} \to H\big(u^{(u)}, u^{(\eta(u))}\big)$$

Also

$$\mathfrak{A}_\eta \& F^{(M)} \to H\big(u, u^{(\eta(0))}\big)$$

Hence for each u some formula of form

$$\mathfrak{A}_\eta \& F^{(M)} \to H\big(u^{(u)}, u^{(\eta(u))}\big)$$

is provable.

Also if $M' \geqslant M$ and $M' \geqslant m$ and $m \neq \eta(u)$

then $\mathfrak{A}_\eta \& F^{(M')} \to G\big(u^{(\eta(u))}, u^{(u)}\big) \vee G\big(u^{(u)}, u^{(\eta(u))}\big)$

and $\mathfrak{A}_\eta \& F^{(M')} \to \Big[\Big\{ G\big(u^{(\eta(u))}, u^{(u)}\big) \vee G\big(u^{(u)}, u^{(\eta(u))}\big) \& H\big(u^{(u)}, u^{(\eta(u))}\big) \Big\} \to \big(-H(u^{(u)}, u^{(u)})\big) \Big]$

Whence

$$\mathfrak{A}_\eta \& F^{(M')} \to \big(-H(u^{(u)}, u^{(u)})\big)$$

Proof of (ii).

Let $H(x,y)$ mean "$\eta(x)=y$", and let $K(x,y,z)$ mean "$\varphi(x,y)=z$". $\mathfrak{A}_\varphi$ is the axiom for $\varphi(x,y)$. We take $\mathfrak{A}_\eta$ to be

$$\mathfrak{A}_\varphi \,\&\, P \,\&\, \big(F(x,y)\to G(x,y)\big)\,\&\,\big(G(x,y)\,\&\,G(y,z)\to G(x,z)\big)$$
$$\&\,\big(F^{(r)}\to H(a,a^{(r)})\big)\,\&\,\big(F(\eta\omega)\,\&\,H(\omega,x)\,\&\,K(\omega,x,z)\to H(\omega,z)\big)$$
$$\&\,\big[H(\omega,z)\,\&\,G(z,t)\vee G(t,z)\to(-H(\omega,t))\big]$$

I shall not give the proof of consistency of $\mathfrak{A}_\eta$. Such a proof may be constructed by the methods used in Hilbert and Bernays Grundlagen der Mathematik (Berlin 1934), p.209 et seq. The consistency is also clear from the meaning.

Suppose that for some n, N we have shewn

$$\mathfrak{A}_\eta \,\&\, F^{(N)}\to H\big(a^{(n-1)},a^{(\eta(n-1))}\big)$$

then for some M

$$\mathfrak{A}_\varphi \,\&\, F^{(M)}\to K\big(a^{(n)},a^{(\eta(n-1))},a^{(\eta(n))}\big)$$
$$\mathfrak{A}_\eta \,\&\, F^{(M)}\to F\big(a^{(n-1)},a^{(n)}\big)\,\&\,H\big(a^{(n-1)},a^{(\eta(n-1))}\big)$$
$$\&\,K\big(a^{(n)},a^{(\eta(n-1))},a^{(\eta(n))}\big)$$

A Note on Normal Numbers

Although it is known that almost all numbers are normal [1] no example of a normal number has ever been given . I propose to shew how normal numbers may be constructed and to prove that almost all numbers are normal constructively

Consider the R -figure integers in the scale of t ($t \geqslant 2$). If γ is any sequence of figures in that scale we denote by $N(t, \gamma, n, R)$ the number of these in which γ occurs exactly n times. Then it can be proved without difficulty that

$$\frac{\sum_{n=2}^{R} n\, N(t, \gamma, n, R)}{\sum_{n=1}^{R} N(t, \gamma, n, R)} = \frac{R - r + 1}{R} t^{-r}$$

where $l(\gamma) = r$ is the lenght of the sequence γ : it is also possible [2] to prove that

$$\sum_{|n - Rt^{-r}| > K} N(t, \gamma, n, R) < 2 t^{R} e^{-K^2 t^r / 4R} \qquad (1)$$

provided $\frac{Kt^r}{R} < \cdot 3$

Let α be a real number and $S(\alpha, t, \gamma, R)$ the number of occurrences of γ in the first R figures after the dedimal point in the expression of α in the scale of t . α is said to be <u>normal</u> if

$$R^{-1} S(\alpha, t, \gamma, R) \to t^{-r}$$

as $R \to \infty$ for each γ, t

where $r = l(\gamma)$.

Now consider sums of a finite number of open intervals with rational end points. These can be enumerated constructively. We take a particular constructive enumeration: let E_n be the n th

set of intervals in the enumeration. Then we have

Theorem 1

We can find a constructive[3] function $c(K,n)$ of two integral variables, such that

$$E_{c(K,n+1)} \subseteq E_{c(K,n)}$$

and $$m\, E_{c(K,n)} > 1 - \frac{1}{K} \quad \text{for each } K, n$$

and $E_{(K)} = \prod_{n=2}^{\infty} E_{c(K,n)}$ consists entirely of normal numbers for each K.

Let $B(\Delta, \gamma, t, R)$ be the set of numbers α, $0<\alpha<1$, for which

$$\left| S(\alpha, t, \gamma, R) - R t^{-r} \right| < \frac{R}{\Delta t^{r}} \qquad \left(K = \frac{R}{\Delta t^{r}}\right) \quad (2)$$

$\Delta = \frac{1}{K t^{r}}$

then by (1)

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

if $\Delta < .3$.

Let $A(\Delta, T, L, R)$ be the set of those α for which (2) holds whenever $2 \le t \le T$ and $l(\gamma) \le L$ i.e.

$$A(\Delta, T, L, R) = \prod_{t=2}^{T} \prod_{l(\gamma)\le L} B(\Delta, \gamma, t, R)$$

The number of factors in the product is at most T^{L+1} so that

$$m\, A(\Delta, T, L, R) > 1 - T^{L+1} e^{-R t^{1-r}/4\Delta^2} e^{-\frac{R T^{1-r}}{4\Delta^2}}$$

Let

$$A_k = A\left([k^{1/4}], [e^{\sqrt{\log k}}], \left[\sqrt{\log k} - 1\right], k \right)$$

$$A_k = A\left(k, [e^{\sqrt{\log k}}], \left[\sqrt{\log k} - 1\right], k^4 \right)$$

then if $k \geqslant 1000$ we shall have $m A_k > 1 - k e^{-\frac{1}{2}k^{1/2}} > 1 - \frac{1}{k(k-1)}$.

$C(K,u)$ ($K \geq 1000$) is to be defined as follows

$C(K,0)$ is $(0,1)$

$C(K,u+1)$ is the intersection of an interval $(\beta_u, 1)$, $(0 \leq \beta_u < 1)$ with A_{K+u+1} and $C(K,u)$, β_u being so chosen that the measure of $C(K,u+1)$ is $1 - \frac{1}{K} + \frac{1}{K+u+1}$. This is possible since the measure of $C(K,u)$ is $1 - \frac{1}{K} + \frac{1}{K+u}$ and that of A_{K+u+1} is at least $1 - \frac{1}{(K+u)(K+u+1)}$; consequently the measure of $C(K,u) \cap A_{K+u+1}$ is at least $1 - \frac{1}{K} + \frac{1}{K+u+1}$.

If $K < 1000$ we define $C(K,u)$ to be $C(1000,u)$. $C(K,u)$ is a finite sum of intervals for each K, u. When we remove the boundary points we obtain a set of form $E_{C(K,u)}$ of measure $1 - \frac{1}{K} + \frac{1}{K+u}$ ($K \geqslant 1000$). The intervals of which $E_{C(K,u)}$ is composed may be found by a mechanical process and so the function $C(K,u)$ is constructive. The set $E_{(K)} = \prod_{u=1}^{\infty} E_{C(K,u)}$ consists of normal numbers, for if $\alpha \in E$ then $\alpha \in A_K$, all $K > K$ and $K > 1000$. If γ is a sequence of length r in the scale of t and if k_0 be such that $\left[e^{\sqrt{\log k_0}}\right] > t$ and $\left[\sqrt{\log k_0}\right] > r+1$ then for $k > k_0$

$|S(\beta, t, \gamma, k) - k t^{-r}| < k \left[k^{1/4}\right]^{-1}$ when β is in A_k (by the definition of A_k). Hence $k^{-1} S(\alpha, t, \gamma, k) \to t^{-r}$ as k tends to infinity, i.e. α is normal.

Theorem 2

There is a rule whereby given an integer K and an infinite sequence of figures 0 and 1 (the p th figure in the sequence being $\vartheta(p)$) we can find a normal number $\alpha(K, \vartheta)$ in the interval (0,1) and in such a way that for fixed K these numbers form a set of measure at least $1 - 2/K$, and so that the first n figures of ϑ determine $\alpha(K, \vartheta)$ to within 2^{-n}.

With each integer n we associate an interval of the form $\left(\frac{m_n}{2^n}, \frac{m_n+1}{2^n}\right)$ whose intersection with $E_{c(\vartheta,n)}$ is of positive measure. and given m_n we obtain m_{n+1} as follows. Put

$$m\, E_{c(\vartheta,n)} \cap \left(\frac{m_n}{2^n}, \frac{2m_n+1}{2^{n+1}}\right) = a_{n,m}$$

$$m\, E_{c(\vartheta,n)} \cap \left(\frac{2m_n+1}{2^{n+1}}, \frac{m_n+1}{2^n}\right) = b_{n,m}$$

and let r_n be the smallest m for which either $a_{n,m} < K^{-1}2^{-2n}$ or $b_{n,m} < K^{-1}2^{-2n}$ or both $a_{n,m} > \frac{1}{K(K+n+2)}$ and $b_{n,m} > \frac{1}{K(K+n+2)}$. There exists such an r_n for $a_{n,m}$ and $b_{n,m}$ decrease either to 0 or to some positive number. In the case where $a_{n,r_n} < K^{-1}2^{-2n}$ we put $m_{n+1} = 2m_n + 1$: if $a_{n,r_n} \geq K^{-1}2^{-2n}$ but $b_{n,r_n} < K^{-1}2^{-2n}$ we put $m_{n+1} = 2m_n$, and in the third case we put $m_{n+1} = 2m_n$ or $m_{n+1} = 2m_n + 1$ according as $\vartheta(n) = 0$ or 1. For each n the interval $\left(\frac{m_n}{2^n}, \frac{m_n+2}{2^{n+1}}\right)$ includes normal numbers in positive measure. The intersection of these intervals contains only one number which must be normal.

Now consider the set $A(K,n)$ consisting of all possible intervals $\left(\frac{m_n}{2^n}, \frac{m_n+1}{2^n}\right)$ i.e. the sum of all these intervals as we allow the first n figures of ϑ to run through all possibilities. Then

$$m\, E_{(K)} \cap A(K,n+1) = m\, E_{(K)} \cap A(K,n) - \sum_{m=0}^{2^n-1} m\, E_{(K)} \cap \left(A(K,n) - A(K,n+1)\right) \cap \left(\frac{m}{2^n}, \frac{m+1}{2^n}\right)$$

For

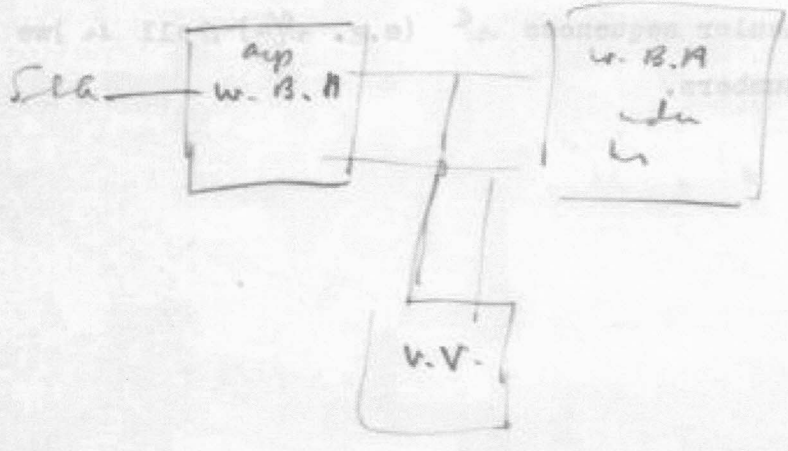
But $$m\left(A(K,n) - A(K,n+1)\right) \cap \left(\frac{m}{2^n}, \frac{m+1}{2^n}\right) < 2^{-2n} K^{-1}$$

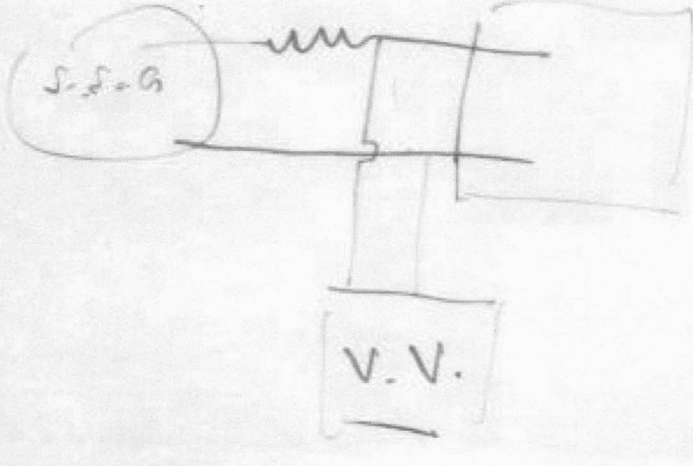
so that

$$m\bar{E}_{(K)} \cap A(K, n+1) > m\bar{E}_{(K)} \cap A(K,n) - 2^{-n-1}K^{-1} > m\bar{E}_{(K)} - K^{-1} > 1 - 2/K$$

The set of alll possible numbers $\alpha(K, \vartheta)$ is therefore of measure at least $1 - 2/K$.

By taking particular sequences ϑ (e.g. $\vartheta(n) = 0$ all n) we obtain particular normal numbers.


W. B. A
W. B. A
V. V.


V. V.

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

Collected Works of A.M. Turing

PURE MATHEMATICS

Edited by

J.L. BRITTON
King's College, London, United Kingdom

with a section on Turing's statistical work by I.J. GOOD
Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

1992
NORTH-HOLLAND
AMSTERDAM · LONDON · NEW YORK · TOKYO

A NOTE ON NORMAL NUMBERS

Although it is known that almost all numbers are normal no example of a normal number has ever been given. I propose to show how normal numbers may be constructed and to prove that almost all numbers are normal constructively. [1]

Consider the R-figure integers in the scale of t, $t \geqslant 2$. If γ is any sequence of figures in that scale we denote by $N(t, \gamma, n, R)$ the number of these in which γ occurs exactly n times. Then it can be proved without difficulty that [2]

$$\left(\sum_{n=1}^{R} n N(t,\gamma,n,R)\right) \Big/ \left(\sum_{n=0}^{R} N(t,\gamma,n,R)\right) = (R-r+1)t^{-r}, \quad [3]$$

where $l(\gamma) = r$ is the length of the sequence γ: it is also possible to prove that

$$\sum_{|n - Rt^{-r}| > k} N(t,\gamma,n,R) < 2t^{R} e^{-k^2 t^r / 4R}, \qquad (1)$$

provided $kt^r/R < 0.3$.

Let α be a real number and $S(\alpha, t, \gamma, R)$ the number of occurrences of γ in the first R figures after the decimal point in the expression of α in the scale of t. α is said to be *normal* if $R^{-1}S(\alpha, t, \gamma, R) \to t^{-r}$ as $R \to \infty$ for each γ, t, where $r = l(\gamma)$.

Now consider sums of a finite number of open intervals with rational end points. These can be enumerated constructively. We take a particular constructive enumeration: let E_n be the nth set of intervals in the enumeration. Then we have the next theorem.

Theorem 1. *We can find a constructive function $c(k, n)$ of two integral variables such that $E_{c(k,n+1)} \leqslant E_{c(k,n)}$ and $m E_{c(k,n)} > 1 - 1/k$ for each k, n and $E(k) = \prod_{n=1}^{\infty} E_{c(k,n)}$ consists entirely of normal numbers for each k.*

Let $B(\Delta, \gamma, t, R)$ be the set of numbers α $(0 < \alpha < 1)$ for which

$$|S(\alpha, t, \gamma, R) - Rt^{-r}| < \frac{R}{\Delta t^r}, \quad \left(K = \frac{R}{\Delta t^r}\right), \quad \Delta = \frac{R}{Kt^r}. \qquad (2)$$

Then by (1)

$$m\, B(\Delta, \gamma, t, R) > 1 - 2e^{-Rt^r/4\Delta^2} \quad \text{if } \Delta^{-1} < 0.3. \qquad [4]$$

Let $A(\Delta, T, L, R)$ be the set of those α for which (2) holds whenever $2 \leqslant t \leqslant T$ and $l(\gamma) \leqslant L$, i.e.,

$$A(\Delta, T, L, R) = \prod_{t=2}^{T} \prod_{l(\gamma) \leqslant L} B(\Delta, \gamma, t, R).$$

The number of factors in the product is at most T^{L+1} so that

[5] $$m A(\Delta, T, L, R) > 1 - T^{L+1} e^{-RT^{-2}/4\Delta^2}.$$

Let

$$A_k = A([k^{1/4}], [e^{\sqrt{\log k}}], [\sqrt{\log k} - 1], k),$$
$$\bar{A}_k = A(k, [e^{\sqrt{\log k}}], [\sqrt{\log k} - 1], k^4).$$

Then, if $k \geqslant 1000$, we shall have

$$m A_k > 1 - k e^{-1/2k^{1/4}} > 1 - 1/k(k-1).$$

$c(k, n)$ $(k \geqslant 1000)$ is to be defined as follows. $c(k, 0)$ is $(0, 1)$. $c(k, n+1)$ is the intersection of an interval $(\beta_n, 1)$ $(0 \leqslant \beta_n < 1)$ with A_{k+n+1} and $c(k, n)$, β_n being chosen so that the measure of $c(k, n+1)$ is $1 - 1/k + (k+n+1)^{-1}$. This is possible since the measure of $c(k, n)$ is $1 - 1/k + 1/(k+n)$ and that of A_{k+n+1} is at least $1 - 1/((k+n)(k+n+1))$. Consequently the measure of
[6] $c(k, n) \cap A_{k+n+1}$ is at least $1 - 1/k + 1/(k+n+1)$. If $k < 1000$ we define $c(k, n)$ to be $c(1000, n)$. $c(k, n)$ is a finite sum of intervals for each k, n. When we remove the boundary points we obtain a set of the form $E_{c(k,n)}$ of measure $1 - 1/k + 1/(k+n)$ $(k \geqslant 1000)$. The intervals of which $E_{c(k,n)}$ is composed may be found by a mechanical process and so the function $c(k, n)$ is constructive. The set $E(k) = \prod_{n=1}^{\infty} E_{c(k,n)}$ consists of normal numbers for if $\alpha \in E(k)$, then $\alpha \in A_k$ (all $k > K$, $k \geqslant 1000$). If γ is a sequence of length r in the scale of t and if k_0 be such that

$$[e^{\sqrt{\log k_0}}] > t \quad \text{and} \quad [\sqrt{\log k_0}] > r + 1,$$

then for $k > k_0$

$$|S(\beta, t, \gamma, k) - k t^{-r}| < k[k^{1/4}]^{-1},$$

where β is in A_k (by the definition of A_k). Hence $k^{-1} S(\alpha, t, \gamma, k) \to t^{-r}$ as k tends to infinity, i.e., α is normal.

Theorem 2. *There is a rule whereby given an integer k and an infinite sequence of figures* 0 *and* 1 *(the pth figure in the sequence being $\theta(p)$) we can find a normal number $\alpha(k, \theta)$ in the interval* $(0, 1)$ *and in such a way*

that for fixed k these numbers form a set of measure at least $1-2/k$ *and so that the first n figures of* θ *determine* $\alpha(k,\theta)$ *to within* 2^{-n}. [7]

With each integer n we associate an interval of the form $(m_n/2^n, (m_n+1)/2^n)$ whose intersection with $E(k)$ is of positive measure, and given m_n we obtain m_{n+1} as follows. Put [8]

$$m E_{c(k,n)} \cap \left(\frac{m_n}{2^n}, \frac{2m_n+1}{2^{n+1}}\right) = a_{n,m}, \qquad [9]$$

$$m E_{c(k,n)} \cap \left(\frac{2m_n+1}{2^{n+1}}, \frac{m_n+1}{2^n}\right) = b_{n,m},$$

and let r_n be the smallest m for which either $a_{n,m} < k^{-1}2^{-2n}$ or $b_{n,m} < k^{-1}2^{-2n}$ or both $a_{n,m} > 1/k(k+n+1)$ and $b_{n,m} > 1/k(k+n+1)$. There exists such an r_n for $a_{n,m}$ and $b_{n,m}$ decrease either to 0 or to some positive number. In the case where $a_{n,r_n} < k^{-1}2^{-2n}$ we put $m_{n+1} = 2m_n + 1$; if $a_{n,r_n} \geqslant k^{-1}2^{-2n}$ but $b_{n,r_n} < k^{-1}2^{-2n}$, we put $m_{n+1} = 2m_n$, and in the third case we put $m_{n+1} = 2m_n$ or $m_{n+1} = 2m_n + 1$ according as $\theta(n) = 0$ or 1.

For each n the interval $(m_n/2^n, (m_n+1)/2^{n+1})$ includes normal numbers in positive measure. The intersection of these intervals contains only one number which must be normal. [10]

Now consider the set $A(k,n)$ consisting of all possible intervals $(m_n/2^n, (m_n+1)/2^n)$, i.e., the sum of all these intervals as we allow the first n figures of θ to run through all possibilities. Then

$$m E(k) \cap A(k,n+1) = m E(k) \cap A(k,n) - \sum_{m=0}^{2^n-1} m E(k) \cap (A(k,n) - A(k,n+1)) \cap \left(\frac{m}{2^n}, \frac{m+1}{2^n}\right).$$

But

$$m (A(k,n) - A(k,n+1)) \cap \left(\frac{m}{2^n}, \frac{m+1}{2^n}\right) < 2^{-2n}k^{-1},$$

so that

$$m E(k) \cap A(k,n+1) > m E(k) \cap A(k,n) - 2^{-n-1}k^{-1} > m E(k) - k^{-1} > 1 - \frac{2}{k}. \qquad [11]$$

The set of all possible numbers $\alpha(K,\theta)$ is therefore of measure at least $1-2/k$.

By taking particular sequences θ (e.g., $\theta(n) = 0$ for all n) we obtain particular normal numbers. [12]

arbitrary matrix with integral entries whether or not it is a product with each factor one of the given matrices (MARKOV (1951)).

[16] A specialization of this is the following. As mentioned in note [14], there is a finitely presented group whose word problem is unsolvable. This raises the question: is there an algorithm which will decide of a given finite presentation whether or not its word problem is unsolvable? That no such algorithm exists was shown by ADYAN (1957).

[17] Markov has shown that there is no algorithm for deciding of two 4-manifolds whether or not they are homeomorphic. The proof is by reduction to the unsolvability of the isomorphism problem for groups, that is, to the result that there is no algorithm for deciding of any two finite presentations of groups whether or not the groups are isomorphic (ADYAN (1957); MARKOV (1958); RABIN (1958)).

[18] (a) Hemion has shown that the decision problem for knots is solvable and that the homeomorphism problem for a large class of 3-manifolds is solvable (HEMION (1979)).

(b) Another decision problem which Turing might have included is Hilbert's 10th problem: is there an algorithm which decides for any polynomial $p(x_1, \ldots, x_n)$ over the integers whether or not there exist integers $a_1, \ldots, a_n$ such that $p(a_1, \ldots, a_n) = 0$?

This was shown to be unsolvable by MATIJASEVIČ (1970); his proof was based on the work of DAVIS, PUTNAM and ROBINSON (1961).

A Note on Normal Numbers

Notes

[1] In fact, CHAMPERNOWNE (1933) gave an example of a normal number.

[2] More accurately, γ is $u_1, \ldots, u_r$ where $0 \leqslant u_i < t$, $i = 1, \ldots, r$, and $u_1, \ldots, u_r$ are all different.

[3] The second summation should be from 0 to R. The right-hand side should be $(R - r + 1)t^{-r}$; however this formula is not needed in the sequel.

[4] Fix R, T and r and simplify the notation to $N(\gamma, n), S(\alpha, \gamma)$. Let J be the set of all R-figure integers in scale t and for $a \in (0, 1)$ let $\varphi(a) \in J$ be the first R figures after the decimal point in the expansion of a in scale t.

If $\xi \in J$, let $S'(\xi, \gamma)$ be the number of occurrences of γ in ξ. We have $N(\gamma, n)$ = number of elements of J in which γ occurs exactly n times. Now

$$\text{number of pairs } (\xi, \gamma) \text{ such that } S'(\xi, \gamma) = x$$
$$= \sum_{\gamma} (\text{number of pairs } (\xi, \gamma) \text{ such that } S'(\xi, \gamma) = x)$$

$$= \sum_{\gamma} (\text{number of } \xi \text{ such that } \gamma \text{ occurs } x \text{ times in } \xi)$$

$$= \sum_{\gamma} N(\gamma, x).$$

Let $\Delta < 0.3$. The number of pairs such that $|S'(\xi, \gamma) - Rt^{-r}| > R/t^r\Delta$ equals $\sum N(\gamma, x)$ over γ and x where x satisfies $|x - Rt^{-r}| > R/t^r\Delta$. This sum is less than $2t^R \exp(-k^2t^r/4R)$. The set of all a such that $\varphi(a)$ is a given ξ is an interval of length t^{-R}. Also, $S(a, \gamma) = S'(\varphi(a), \gamma)$. Hence the measure of the set of all a in (0, 1) such that $|S(a, \gamma) - Rt^{-r}| > R/t^r\Delta$ is less than $2\exp(-k^2t^r/4R)$, that is, $2\exp(-Rt^{-r}/4\Delta^2)$. The required inequality follows.

[5] Replace T^{L+1} by $2T^{L+1}$.

[6] A_{k+n} is a finite sum (i.e., union) of intervals since each $B(\Delta, \gamma, t, R)$ is; this is essentially because in the notation of note [4], $\varphi^{-1}(\xi)$ is an interval for each ξ in J.

Moreover the end-points of these intervals are rational.

T2

[7] The proof of this theorem that is given is certainly inadequate. Indeed I suspect that the theorem is false.

In the theorem let $\theta(p) = 0$ for all p or, more generally, let θ be recursive. Then in the notation of the proof we can recursively enumerate $m_1, m_2, \ldots$, so the intervals $I_n = (m_n/2^n, (m_n + 1)/2^n)$ are recursively enumerable. Assuming that the intersection of these intervals is a single point x and noting that $m(I_n) = 2^{-n}$ we see that x is not Martin-Löf random (see CHAITIN (1987)).

[8] There seems to be a serious gap in the argument giving the construction of these intervals. It is trivial to find such intervals nonconstructively but we have to proceed constructively.

Note that if m and n are given we can calculate $a_{n,m}$ and $b_{n,m}$ (see note [9]). Suppressing n, we have that (a_m) is a decreasing sequence with limit $a \geqslant 0$ (unknown) and similarly $b_m \to b$ (unknown). Let $c = a + b$; then $c > 0$. Assume that $0 < c' \leqslant c \leqslant c''$, where c' and c'' are known and $c'' < 2c'$. Choose x and y such that $0 < x < c'$, $0 < y < c'$, $c'' < x + y$. Then some $a_m < x$ or some $b_m < x$ or for some m both $a_m < y$ and $b_m < y$; otherwise, for all m, $a_m + b_m \geqslant x + y$ so $c \geqslant x + y > c''$.

In the first case $a < x$ so $b > c - x \geqslant c' - x > 0$; similarly in the third case $a < y$, $b < y$ so $a > c' - y > 0$ and $b > c' - y > 0$. (This analysis shows that perhaps '$> 1/k(k + n + 1)$' should be '$< 1/k(k + n + 1)$'.)

In the first case we choose the right half-interval, in the second case the left one and in the third case we make a choice determined by θ.

If something of the form above is what the author had in mind for ob

taining I_{n+1} from I_n, I see no way of defining x, y, c', c'' as functions of n which would allow the construction of all the intervals I_n.

[9] The symbol 'm' on the left clearly denotes the measure of the intersection of the two sets. The right-hand side depends on n and another variable, unfortunately called m, so evidently $c(k,n)$ should be $c(k,m)$.

[10] What is to exclude the case when, from some point onwards, always the left half-interval is chosen? In this case the intersection of the intervals is empty. (However, if the intersection is empty, then the point in common to all the corresponding closed intervals is rational.)

Assuming the intersection is nonempty, why is the number normal? (Consider the analogous situation: $I_n=(\frac{1}{2}-\frac{1}{2}/n, \frac{1}{2}+\frac{1}{2}/n)$, $E=(0,\frac{1}{2})\cup(\frac{1}{2},1)$. Then $I_n\cap E$ has positive measure, $\bigcap I_n=\{\frac{1}{2}\}$ but $\frac{1}{2}$ is not in E.)

[11] Replace 2^{-n-1} by 2^{-n} (and replace $1-2/k$ by $1-3/k$)?

[12] There are three reference numbers in the text but there are no corresponding references.

For further information on normal numbers, see KUIPERS and NIEDERREITER (1974).

The Word Problem in Compact Groups

Notes

[1] Let L be the first-order language with equality having one binary relation symbol $<$ and the binary function symbols $-$ and $\times$. A sentence A of L is said to be *true for* $\mathbb{R}$ if it is true for the real numbers when $=, <, -, \times$ have their usual meanings. Tarski's theorem may be expressed by saying that there is an algorithm which given any sentence A determines whether or not A is true for $\mathbb{R}$.

Note for later that we can express $y=0$ in L as $y-y=y$. Also,

$$yy=y \quad \text{and not } y-y=y$$

expresses $y=1$, $z=x-((x-x)-y)$ expresses $z=x+y$, and the formula $zy=x$ and not $y-y=y$ expresses $z=xy^{-1}$.

[2] For procedure (a) to make sense, the group G must be given in terms of generators and defining relations, the set of generators must be recursively enumerable and the set of defining relations must also be recursively enumerable. In particular G must be countable; this rules out, for example, the unitary group $U(n)$.

[3] Before 'matrices' insert 'complex nonsingular'.

'order r' means of course that the matrices are $r\times r$.

[4] This means that, if the statement in quotation marks is denoted by S_r, then there is a sentence A_r in the first-order language in note [1] such that