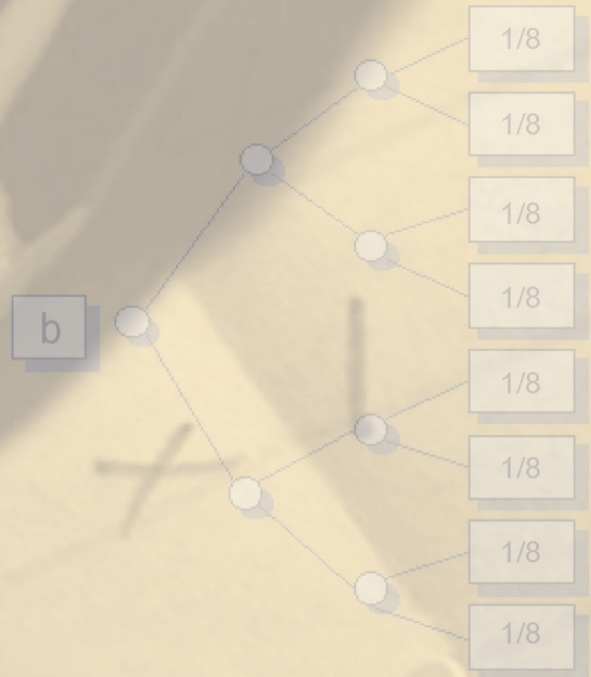


MAURICIO CORREA VILLA

Fundamentos de la teoría de la información



FUNDAMENTOS DE LA TEORÍA DE LA INFORMACIÓN

MAURICIO CORREA VILLA





INSTITUTO TECNOLÓGICO METROPOLITANO
Institución Universitaria

FUNDAMENTOS DE LA TEORÍA DE LA INFORMACIÓN

© Mauricio Correa Villa

© Instituto Tecnológico Metropolitano

1a. Edición: diciembre de 2008

ISBN: 978-958-8351-52-0

Dirección editorial
Oficina de Comunicaciones y Publicaciones

Diagramación y montaje
L. Vieco e Hijos Ltda.

Impreso y hecho en Medellín, Colombia

*Las opiniones, originalidad y citas del texto son responsabilidad del autor.
El Instituto salva cualquier obligación derivada del libro que se publica. Por lo
tanto, ella recaerá única y exclusivamente en el autor.*

Instituto Tecnológico Metropolitano
Calle 73 No. 76A 354
Tel.: (574) 440 51 00
Fax: 440 51 01
www.itm.edu.co
Medellín - Colombia

CONTENIDO

INTRODUCCIÓN	11
1. CONCEPTOS PRELIMINARES SOBRE LA TEORÍA DE LA INFORMACIÓN.....	13
1.1 La teoría de la información y la comunicación	13
1.2 Modelo de comunicación	13
1.3 Información.....	16
1.4 Principios de la medición de información	18
1.5 Unidad de información	19
1.6 Redundancia	22
1.7 Capacidad del canal	24
2. LA TRANSMISIÓN DE LA INFORMACIÓN.....	27
2.1 Qué es teoría de la información	27
2.2 Codificación de la información	27
2.3 Un problema en la transmisión de información.....	29
3. LA INFORMACIÓN Y SUS FUENTES.....	33
3.1 Definición de información.....	33
3.2 Fuente de información de memoria nula	35
3.3 Propiedad fundamental de la entropía	36
3.4 Extensiones de una fuente de memoria nula.....	38
3.5 Fuente de información de Markov.....	39
4. PROPIEDADES DE LOS CÓDIGOS	49
4.1 Definición	49
4.2 Códigos unívocamente decodificables	49
4.3 Inecuación de kraft	54

5.	CODIFICACIÓN	61
5.1	Longitud media de un código	61
5.2	Código compacto	61
5.3	Codificación Huffman	65
5.4	Construcción de códigos compactos binarios	67
5.5	Códigos compactos r-arios.....	70
5.6	Rendimiento y redundancia de un código	72
6.	ANÁLISIS DE CANALES	77
6.1	Canales de información	77
6.2	Probabilidades asociadas a un canal	80
6.3	Entropías a priori y a posteriori	82
6.4	Equivocación de un canal	83
6.5	Información mutua.....	83
6.6	Propiedades de la información mutua.....	84
6.7	Canales sin ruido y canales determinantes	84
6.8	Canales en serie	88
6.9	Canales reducidos y reducciones suficientes.....	89
	APÉNDICE. LA FUNCIÓN LOGARÍTMICA	95
	BIBLIOGRAFÍA	97

PREFACIO

Este material pretende aproximar al estudioso a las bases de la Teoría Matemática de la Información. El lector que desee ampliar el conocimiento en algunos de los temas tratados o de responder a sus propias necesidades con mayor rigurosidad, deberá consultar obras más especializadas, algunas de las cuales se proponen en la bibliografía.

El texto se deriva de un estudio amplio sobre la obra de Norman Abramson, *Teoría de la Información y Codificación*, complementada con apartes de la *Introducción a la psicología de la comunicación*, material editado por la Universidad Católica de Chile. Los aportes del autor se entremezclan como un conjunto de notas y ejercicios preparados para un curso semestral con estudiantes de Ingeniería Informática. Ninguna de las teorías expuestas son originales, se han dado sí, formas y estilos propios heredados de la experiencia docente en los cursos de Teoría de la Información y algunos de Matemáticas básicas y operativas.

Se sugiere al lector una revisión previa a lo concerniente sobre la función logarítmica y sus propiedades, resolución de sistemas de ecuaciones lineales, definición de probabilidad, en especial, los conceptos sobre probabilidades conjunta, marginal, y condicional, eventos y al teorema de Bayes.

Damos los agradecimientos al ingeniero Fabio Suárez, docente de área de telecomunicaciones del ITM, por sus aportes y revisión al material.

INTRODUCCIÓN

La primera mitad del siglo pasado se caracterizó por un creciente avance de los medios de comunicación, y por la forma en el procesamiento y transmisión de la información. Así pues, se desarrolla el primer modelo científico del proceso de comunicación conocido como la Teoría de la Información o Teoría Matemática de la Comunicación. Específicamente, se desarrolla en el área de la telegrafía donde surge la necesidad de determinar, con la máxima precisión, la capacidad de los diferentes sistemas de comunicación para transmitir información*.

La primera formulación de las leyes matemáticas que gobiernan dicho sistema fue realizada por Hartley (1928) y sus ideas son consideradas actualmente como la génesis de la Teoría de la Información. Posteriormente, Shannon y Weaver (1949) desarrollaron los principios definitivos de esta teoría. Su trabajo se centró en algunos de los siguientes problemas que surgen en los sistemas destinados a manipular información: cómo hablar los mejores métodos para utilizar los diversos sistemas de comunicación; cómo establecer el mejor método para separar las señales del ruido, y cómo determinar los límites posibles de un canal.

El concepto de comunicación en el contexto de la Teoría de la Información es empleado en un sentido muy amplio en el que “quedan incluidos todos los procedimientos mediante los cuales una mente puede influir en otra”. De esta manera, se consideran todas las formas que el hombre utiliza para transmitir sus ideas: la palabra hablada, escrita o transmitida (teléfono, radio, telégrafo, etc.), los gestos, la música, las imágenes, los movimientos.

En el proceso de comunicación es posible distinguir, por lo menos, tres niveles de análisis diferentes: el técnico, el semántico y el pragmático. En el nivel técnico se analizan aquellos problemas que surgen en torno a la fidelidad con que la información puede ser transmitida desde

* [López]

el emisor hasta el receptor. En el semántico se estudia todo aquello que se refiera al significado del mensaje y su interpretación. Por último, en el nivel pragmático se analizan los efectos conductuales de la comunicación, la influencia o efectividad del mensaje en tanto da lugar a una conducta. Es importante destacar que la Teoría de la Información se desarrolla como una respuesta a los problemas técnicos del proceso de comunicación, aun cuando sus principios puedan aplicarse en otros contextos.

CAPÍTULO 1. CONCEPTOS PRELIMINARES SOBRE LA TEORÍA DE LA INFORMACIÓN

1.1 LA TEORÍA DE LA INFORMACIÓN Y LA COMUNICACIÓN

La Teoría de la Información tiene sus inicios con la invención del telégrafo y con la definición del código Morse. Samuel Morse trabajó sobre dicho código considerando apenas tres combinaciones posibles: el punto (como resultado de una descarga eléctrica), el trazo (resultado de una corriente eléctrica aplicada continuamente durante un lapso de tiempo) y la ausencia de corriente, que daba como resultado espacios en blanco entre dos señales gráficas. Con esas posibilidades (punto, trazo y espacio), Morse desarrolló un concepto que sería la génesis de la Teoría de la Información: la relación entre ocurrencia y tamaño de los caracteres. Morse verificó que la letra de mayor frecuencia del alfabeto inglés es la letra e y de esa forma la señalizó apenas con un punto (.). Las letras menos frecuentes eran señalizadas por combinaciones mayores, por ejemplo: una letra menos usual como la letra v en la lengua inglesa fue caracterizada por la señal (...-), al mismo tiempo que la coma, más rara todavía, fue caracterizada por el símbolo (--...-). Así, cuanto más recurrente fuese un símbolo necesario para la comunicación, menor era la señal y, por lo tanto, cuanto menor recurrente, mayor era la señal. Morse objetivaba así la economía del tiempo y la energía en la transmisión de datos. Este concepto sufrió varias evoluciones a lo largo del tiempo.

1.2 MODELO DE COMUNICACIÓN

El modelo comunicacional desarrollado por Shannon y Weaver se basa en un sistema de comunicación general que puede ser representado de la siguiente manera:

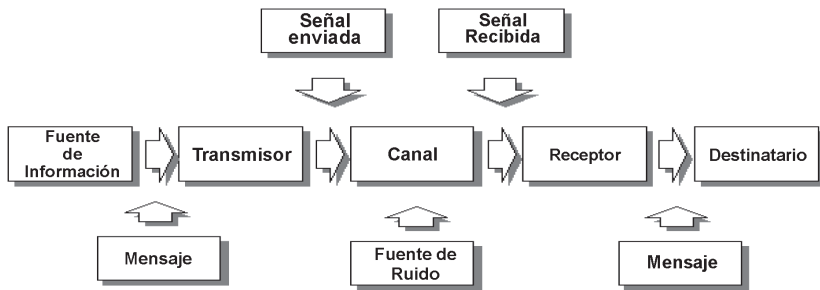


FIGURA 1.1 MODELO COMUNICACIONAL

- Fuente de Información: Selecciona el mensaje deseado de un conjunto de mensajes posibles. Contiene un conjunto de mensajes que pueden seleccionarse para ser enviados por medio de un transmisor
- Transmisor: Transforma o codifica esta información en una forma apropiada al canal. Convierte el mensaje en una señal que será enviada por un canal de comunicación hasta llegar al receptor, que se encargará de convertirla nuevamente en un mensaje que recibirá el destinatario
- Señal: Mensaje codificado por el transmisor
- Canal: Medio a través del cual las señales son transmitidas al punto de recepción
- Fuente de Ruido: Conjunto de distorsiones o adiciones no deseadas por la fuente de información que afectan a la señal. Pueden consistir en distorsiones del sonido (radio, teléfono), distorsiones de la imagen (T.V.), errores de transmisión (telégrafo), etc.
- Receptor: Decodifica o vuelve a transformar la señal transmitida en el mensaje original o en una aproximación de este haciéndolo llegar a su destino

Por ejemplo, cuando una persona comunica algo a otra, la fuente de información es su cerebro, el transmisor son sus cuerdas vocales que generan la señal auditiva que será transmitida por medio del aire

(canal de comunicación) hasta llegar a los oídos (receptor) de la otra persona y convertida nuevamente por medio de su cerebro en el mensaje original.

Expresado de otra forma, cuando yo hablo con usted, actúan los siguientes elementos o partes:

Mi cerebro: Fuente de información.

Su cerebro: Destinatario.

Mi sistema vocal: Transmisor.

Su oído con su octavo par de nervios craneanos: Receptor.

Este sistema de comunicación (modelo comunicacional) es lo suficientemente amplio como para incluir los diferentes contextos en que se da la comunicación (conversación, T.V., danza, etc.). Tomemos como ejemplo lo que ocurre en el caso de la radio. La fuente de información corresponde a la persona que habla por el micrófono. El mensaje son las palabras y sonidos que esta persona emite. El micrófono y el resto del equipo electrónico constituyen el transmisor que transforma este mensaje en ondas electromagnéticas, las cuales corresponden a la señal. El espacio que existe entre las antenas transmisoras y receptoras es el canal, mientras que lo que altera la señal original constituye la fuente de ruido. El aparato de radio de cada hogar es el receptor y el sonido que éste emite corresponde al mensaje recobrado. Las personas que escuchan este mensaje radial son los destinatarios.

También podemos ejemplificar esto mediante el texto que usted lee en este momento. En este caso, nuestros cerebros son la fuente de información y nuestros pensamientos, el mensaje. La máquina a través del procesador de texto constituye el transmisor que transforma nuestros pensamientos en lenguaje escrito, el cual corresponde a la señal. El papel es el canal y cualquier error de tipeo o puntuación, manchas, espacios en blanco, etc., constituyen la fuente de ruido. Por último, usted que está leyendo este ejemplo es a la vez el receptor y destinatario, que a través de la lectura recobra el mensaje por nosotros enviado.

Es importante considerar que el problema del significado del mensaje no es relevante en este contexto. El interés principal de la Teoría de

la Información lo constituye todo aquello relacionado con la capacidad y fidelidad para transmitir información de los diferentes sistemas de comunicación. En el ejemplo anterior, el mensaje podría haber consistido en una secuencia de letras carentes de todo significado e igualmente el problema de cuánta información es transmitida estaría presente. En un sentido amplio, la Teoría de la Información trata acerca de la cantidad de información que es transmitida por la fuente al receptor al enviar un determinado mensaje, sin considerar el significado o propósito de dicho mensaje. No interesa tanto la pregunta: “¿Qué tipo de información?” sino más bien, “¿Cuánta información?” es la que transmite la fuente.

1.3 INFORMACIÓN

Antes de analizar lo que se refiere a la capacidad y fidelidad de un canal determinado para transmitir información, es necesario precisar los alcances de este último concepto. El concepto de información es definido en términos estrictamente estadísticos, bajo el supuesto que puede ser tratado de manera semejante a como son tratadas las cantidades físicas como la masa y la energía. La palabra “información” no está relacionada con lo que decimos, sino más bien, con lo que podríamos decir. El concepto de información se relaciona con la libertad de elección que tenemos para seleccionar un mensaje determinado de un conjunto de posibles mensajes. Si nos encontramos en una situación en la que tenemos que elegir entre dos únicos mensajes posibles, se dice, de un modo arbitrario, que la información correspondiente a esta situación es la unidad. La Teoría de la Información, entonces, conceptualiza el término información como el grado de libertad de una fuente para elegir un mensaje de un conjunto de posibles mensajes.

El concepto de información supone la existencia de duda o incertidumbre. La incertidumbre implica que existen diferentes alternativas que deberán ser elegidas, seleccionadas o discriminadas. Las alternativas se refieren a cualquier conjunto de signos contruidos para comunicarse, sean estos letras, palabras, números, ondas, etc. En este contexto, las señales contienen información en virtud de su potencial para hacer

elecciones. Estas señales operan sobre las alternativas que conforman la incertidumbre del receptor y proporcionan el poder para seleccionar o discriminar entre algunas de estas alternativas.

Se asume que en los dos extremos del canal de comunicación –fuente y receptor– se maneja el mismo código o conjunto de signos. La función de la fuente de información será seleccionar sucesivamente aquellas señales que constituyen el mensaje y luego transmitir las al receptor mediante un determinado canal.

Existen diversos tipos de situaciones de elección. Las más sencillas son aquellas en que la fuente escoge entre un número de mensajes concretos. Por ejemplo, elegir una entre varias postales para enviarle a un amigo. Otras situaciones más complejas son aquellas en que la fuente realiza una serie de elecciones sucesivas de un conjunto de símbolos elementales tales como letras o palabras. En este caso, el mensaje estará constituido por la sucesión de símbolos elegidos. El ejemplo más típico aquí es el del lenguaje.

Al medir cuánta información proporciona la fuente al receptor al enviar un mensaje, se parte del supuesto que cada elección está asociada a cierta probabilidad, siendo algunos mensajes más probables que otros. Uno de los objetivos de esta teoría es determinar la cantidad de información que proporciona un mensaje, la cual puede ser calculada a partir de su probabilidad de ser enviada.

El tipo de elección más simple es el que existe entre dos posibilidades, en que cada una tiene una probabilidad de $1/2$ (0,5). Por ejemplo, al tirar una moneda al aire ambas posibilidades –cara y sello– tienen la misma probabilidad de salir. El caso del lenguaje e idioma es diferente. En éstos la elección de los símbolos que formarán el mensaje dependerá de las elecciones anteriores. Por ejemplo, si en el idioma español el último símbolo elegido es “un”, la probabilidad que la siguiente palabra sea un verbo es bastante menor que la probabilidad que sea un sustantivo o un adjetivo. Asimismo, la probabilidad que a continuación de las siguientes tres palabras “el esquema siguiente” aparezca el verbo “representa” es bastante mayor que la probabilidad que aparezca “pera”. Incluso se ha

comprobado que, en el caso del lenguaje, es posible seleccionar aleatoriamente letras que luego son ordenadas según sus probabilidades de ocurrencia y éstas tienden a originar palabras dotadas de sentido.

1.4 PRINCIPIOS DE LA MEDICIÓN DE INFORMACIÓN

De acuerdo con estas consideraciones probabilísticas es posible establecer un primer principio de la medición de información. Éste establece que mientras más probable sea un mensaje menos información proporcionará. Esto puede expresarse de la siguiente manera:

$$I(x_i) > I(x_k) \text{ si y sólo si } p(x_i) < p(x_k)$$

Donde:

$I(x_i)$: cantidad de información proporcionada por x_i y $p(x_i)$:
probabilidad de x_i

De acuerdo con este principio, es la probabilidad que tiene un mensaje de ser enviado y no su contenido, lo que determina su valor informativo. El contenido sólo es importante en la medida que afecta la probabilidad. La cantidad de información que proporciona un mensaje varía de un contexto a otro, porque la probabilidad de enviar un mensaje varía de un contexto a otro.

Un segundo principio que guarda relación con las elecciones sucesivas establece que si son seleccionados los mensajes X e Y, la cantidad de información proporcionada por ambos mensajes será igual a la cantidad de información proporcionada por X más la cantidad de información proporcionada por Y, dado que X ya ha sido seleccionada. Esto puede ser expresado así:

$$I(x_i \text{ e } y_j) = F(p(x_i)) + F(p(y_j/x_i))$$

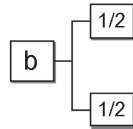
Donde:

$I(x_i \text{ e } y_j)$: cantidad de información proporcionada por los mensajes x_i e y_j F : función $p(x_i)$: probabilidad de x_i $p(y_j/x_i)$: probabilidad de y_j dado que x_i ha sido seleccionado.

1.5 UNIDAD DE INFORMACIÓN

Una vez que hemos seleccionado el mensaje expresado en un lenguaje determinado es posible transcribirlo a un código de tipo binario. Éste consta de sólo dos tipos de señales que indican Si o No, y que generalmente se codifican como 1 ó 0. La cantidad de información proporcionada por cada elección entre dos alternativas posibles constituye la unidad básica de información, y se denomina dígito binario, o abreviadamente bit.

La elección existente al tener un bit de información puede ser esquematizada de la siguiente manera:



En la elección (b) tanto la línea superior como la inferior, es decir ambas posibilidades, pueden ser elegidas con la misma probabilidad de $1/2$.

Si existen N posibilidades, todas igualmente probables, la cantidad de información será igual a $\text{Log}_2 N$. Es, entonces, el $\text{Log}_2 N$ la función matemática que nos indicará la cantidad de bits de información de una situación determinada. Esto puede esquematizarse de la siguiente manera:

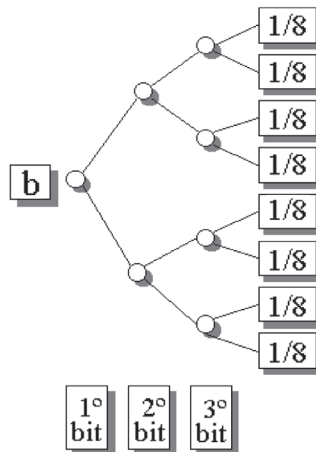


FIGURA 1.2

La Figura 1.2 nos muestra una situación con ocho posibilidades, cada una con una misma probabilidad de $1/8$. Para poder determinar una posibilidad específica de estas ocho, la elección requiere como mínimo tres etapas, cada una de las cuales arroja un bit de información. El primer bit corresponde a la elección entre las primeras cuatro o segundas cuatro posibilidades. El segundo bit corresponde al primer o segundo par de las cuatro posibilidades ya elegidas. El último bit determina el primer o segundo miembro del par y especifica la posibilidad elegida. Como vemos, el primero de bits que se requieren en esta situación para determinar una posibilidad específica es de tres, lo que corresponde al $\text{Log}_2 8$.

Veamos ahora algunos ejemplos de lo recién expuesto:

TABLA 1.1

Signo	Elecciones		
	1°	2°	3°
A	1	1	1
B	1	1	0
C	1	0	1
D	1	0	0
E	0	1	1
F	0	1	0
G	0	0	1
H	0	0	0

La Tabla 1.1 nos muestra un alfabeto compuesto por sólo ocho signos. Pensemos que una fuente de información selecciona un signo y de alguna manera se lo señala al receptor. La pregunta sería entonces, ¿cuánta información deberá conocer el receptor para identificar correctamente el signo escogido?

Asumamos que a partir de elecciones anteriores sabemos que cada uno de los ocho signos tiene la misma probabilidad de ser seleccionado.

La incertidumbre, entonces, se ha repartido uniformemente sobre nuestro “alfabeto”, o lo que es lo mismo, las probabilidades a priori de los signos son iguales; en este caso $1/8$.

Las señales que llegan al receptor representan instrucciones para seleccionar alternativas. La primera instrucción responde a la pregunta ¿está en la primera mitad del alfabeto, si o no? (en la figura, si = 1 y no = 0). La respuesta nos proporciona un bit de información y reduce el rango de incertidumbre exactamente a la mitad. Luego, una segunda instrucción divide cada mitad nuevamente en la mitad y, una tercera instrucción, otra vez en la mitad. En este caso, bastan tres simples instrucciones Si-No (1-0) para identificar un signo cualquiera de un total de ocho. La letra F, por ejemplo, podría ser identificada de la siguiente manera: 010. La respuesta a nuestra pregunta es, entonces, ¡el receptor deberá obtener tres bits de información para identificar correctamente el signo escogido!

El típico juego de las “Veinte Preguntas” ilustra también algunas de las ideas mencionadas. Este juego consiste en que una persona piensa en un objeto mientras el resto de los jugadores intenta adivinar de qué objeto se trata, haciendo no más de veinte preguntas que sólo pueden ser respondidas Si o No. De acuerdo con la Teoría de la Información, cada pregunta y su respuesta pueden proporcionar desde ninguna información hasta un bit de información ($\log_2 2$), dependiendo de si las probabilidades de obtener resultados Si o No son muy desiguales o casi iguales, respectivamente.

Para obtener la mayor cantidad de información posible los jugadores deberán hacer preguntas que dividan el conjunto de posibles objetos en dos grupos igualmente probables. Por ejemplo, si mediante preguntas previas se ha establecido que se trata de una ciudad específica, una buena pregunta sería “¿Está al sur del río X?”. Así se dividen las ciudades posibles en dos grupos aproximadamente iguales. La segunda pregunta podría ser “¿Está al sur del río Y?”. Y así sucesivamente hasta determinar de qué ciudad se trata. Si fuera posible hacer preguntas que tuvieran la propiedad de subdividir las posibilidades existentes en dos grupos

relativamente iguales, sería posible identificar mediante veinte preguntas un objeto entre aproximadamente un millón de posibilidades. Esta cifra corresponde a los 20 bits que se requieren para identificarla ($\log_2 1.000.000$).

1.6 REDUNDANCIA

No obstante lo anterior, la mayoría de las fuentes de información producen mensajes que no consisten en una única elección entre posibilidades de igual probabilidad, sino en elecciones sucesivas entre posibilidades de probabilidad variable y dependiente. A este tipo de secuencias se les denomina procesos estocásticos. Como ya lo mencionamos, el caso más típico son las letras y palabras que conforman el lenguaje. El escribir en español constituye un proceso de elecciones dependientes. Por ejemplo, al formar una palabra se elige una primera letra de todas las posibles primeras letras con diferentes probabilidades; luego, se elige la segunda letra cuya probabilidad depende de la primera letra seleccionada, y así sucesivamente hasta formar la palabra deseada. Lo mismo ocurre en el caso de las palabras para formar oraciones.

Lo importante aquí es señalar el hecho de que, en la medida que se avanza en la formación de una palabra u oración, el rango de posibles letras o palabras a ser seleccionadas va disminuyendo y la probabilidad de que ciertas letras o palabras específicas sean seleccionadas va aumentando. Dicho de otra forma, tanto la incertidumbre como la información de las últimas letras de una palabra o de las últimas palabras de una oración es menor comparada con las primeras.

La mayoría de los mensajes se constituyen a partir de un número limitado de posibilidades, por ejemplo, sólo 29 letras en el caso de nuestro idioma. Como vimos, la probabilidad de ocurrencia de una de estas posibilidades dentro de un mensaje depende de las posibilidades seleccionadas previamente; por ejemplo, la probabilidad de que ocurra la letra “q” luego de una “p” es 0. Son estos dos hechos los que en conjunto

determinan que todo mensaje contenga cierto grado de redundancia. En otras palabras, la redundancia se refiere a que las posibilidades dentro de un mensaje se repiten, y se repiten de una cierta manera predecible. Mientras mayor sea, entonces, la redundancia de un mensaje, menor será su incertidumbre y menor la información que contenga.

El inglés escrito es un tipo de fuente de información que ha sido ampliamente estudiado. Se ha llegado a determinar que la redundancia de la lengua inglesa esta muy próxima al 50%. Es decir, al escribir inglés aproximadamente la mitad de las letras y palabras que se emplean dependen de la libre elección de quien escribe, mientras que la otra mitad está determinada por la estructura probabilística del idioma.

La redundancia de los idiomas permite que si se pierde una fracción de un mensaje sea posible completarlo en forma muy aproximada al original. Este hecho se puede observar al eliminar varias letras de una oración sin que ello impida al lector completar las omisiones y rehacer la oración. Por ejemplo, en la siguiente frase han sido omitidas las vocales: CMPLT ST FRS.

Otra función importante de la redundancia es que nos permite ahorrar tiempo en la decodificación de los mensajes. Generalmente, no leemos cada una de las letras y palabras que conforman un texto, sino que vamos adivinando lo que viene. En el caso del telégrafo, por ejemplo, podríamos ahorrar tiempo ideando un código poco redundante y transmitiendo el mensaje a través de un canal sin ruido. Sin embargo, cuando el canal utilizado tiene ruido es conveniente no emplear un proceso de codificación que elimine toda la redundancia, pues la redundancia nos ayuda a combatir el ruido. Si se pierde parte del mensaje por el ruido que afecta al canal, la redundancia nos permite rehacer en forma aproximada el mensaje. Por el contrario, la fracción de un mensaje no redundante que se pierde por el ruido es imposible de ser recuperada. La redundancia de los mensajes nos permite, entonces, corregir con facilidad los errores u omisiones que hayan podido ocurrir durante la transmisión.

1.7 CAPACIDAD DEL CANAL

Ahora que ya hemos precisado el concepto de información y los conceptos relacionados con él (incertidumbre, bit, redundancia) podemos volver a plantearnos el problema inicial de definir la capacidad de un canal determinado para transmitir información. Dado un canal con una capacidad de C unidades por segundo que recibe señales de una fuente de información de H unidades por segundo, la pregunta es ¿cuánto es el máximo número de bits por segundo que puede ser transmitido a través de este canal? Por ejemplo, un teletipo consta de 32 símbolos posibles que supondremos son empleados con igual frecuencia. Cada símbolo representa entonces 5 bits ($\log_2 32$) de información. De esta forma, si en ausencia total de ruido podemos enviar N símbolos por segundo a través de este canal, entonces podremos enviar $5N$ bits de información por segundo a través de dicho canal.

Son estas dos cantidades, la tasa de transmisión H por la fuente de información y la capacidad C del canal, las que determinan la efectividad del sistema para transmitir información. Si $H > C$ será ciertamente imposible transmitir toda la información de la fuente, no habrá suficiente espacio disponible. Si $H \leq C$ será posible transmitir la información con eficiencia. La información entonces, puede ser transmitida por el canal solamente si H no es mayor que C .

El teorema fundamental para un canal sin ruido que transmite símbolos discretos afirma que si se emplea un procedimiento adecuado de codificación para el transmisor es posible conseguir que el ritmo medio de transmisión de símbolos por el canal sea muy próximo a C/H . Por muy perfecto que sea el procedimiento de codificación, dicho ritmo nunca podrá ser mayor de C/H .

Sin embargo, el problema de calcular la capacidad del canal se complica por la presencia de ruido. La presencia de ruido durante la transmisión provocará que el mensaje recibido contenga ciertos errores que contribuirán a aumentar la incertidumbre. Recordemos que la información es una medida del grado de libertad de elección que poseemos al momento de seleccionar un mensaje. Cuanto mayor sea la libertad de

elección, mayor será la falta de seguridad en el hecho de que el mensaje enviado sea uno determinado. La incertidumbre será mayor y mayor la cantidad de información posible. De esta forma, si el ruido aumenta la incertidumbre, aumentará la información. Esto parecería indicar que el ruido es beneficioso, puesto que cuando hay ruido, la señal recibida es seleccionada a partir de un mayor conjunto de señales que las deseadas por el emisor. Sin embargo, la incertidumbre originada por la libertad de elección del emisor es una incertidumbre deseable; la incertidumbre debida a errores por la influencia del ruido es una incertidumbre no deseable.

Para extraer la información útil de la señal recibida es necesario suprimir la ambigüedad introducida por el ruido. Para ello se recurre a un factor de corrección matemático. El teorema para la capacidad de un canal con ruido se define como el ritmo máximo a que la información útil (incertidumbre total menos la incertidumbre debida al ruido) puede ser transmitida a través del canal.

BIBLIOGRAFÍA

- ABRAMSON, Norman. Teoría de la Información y Codificación. Madrid: PARANINFO, 1966.
- GRAY, Robert M. Entropy and Information Theory. New York, 1990.
- JOHANSEN, Bertoglio Óscar. Introducción a la Teoría General de Sistemas. México: Limusa, 1991. 167 p.
- LÓPEZ Alejandro; PARADA Andrea; SIMONETTI Franco. “Introducción a la psicología de la comunicación”. Santiago: Universidad Católica de Chile, 1995.
- FERREIRA, Marcelo R. Data Mining basado en la Teoría de la Información. Agosto 2007.
- RIBEIRO, Francisco Carlos. Revista Libertas, 41. El pensamiento de Hayek y la Teoría de la Información. Octubre 2004.



Fundamentos de la teoría de la información

se terminó de imprimir en diciembre de 2008.

Para su elaboración se utilizó papel Bond de Alta Blancura 75 g,
en páginas interiores, y cartulina Propalcote 240 g para la carátula.

Las fuentes tipográficas empleadas son Times New Roman 11 puntos,
en texto corrido, y Myriad Pro 14 puntos en títulos.