Universidad de Granada

Ingeniería Informática

Computación y Sistemas Inteligentes

Relacion Tema 1

Autor: José Antonio Ruiz Millán

Asignatura: Teoria de la Información y de la Codificación

20 de octubre de 2018

1. **¿Qué es la tasa de información? ¿Cómo ayuda la tasa de información a comparar dos fuentes? Supongamos que dos fuentes de información A y B tienen una esperanza de información EI(A)= 0.8 y EI(B)= 0.5. Sin embargo, A tarda 3ms en transmitir un símbolo, mientras que B tarda 1ms. ¿Qué fuente tiene una mayor tasa de información?**

   Llamamos tasa de información al cociente entre la información que una fuente es capaz de enviar respecto al tiempo que se tarda en enviar un símbolo.

   Este término nos permite comparar dos fuente en rendimiento, es decir, nosotros podemos tener dos fuentes que tengan la misma *esperanza de información* y que no podríamos decir sólo basandonos en la *esperanza* cual de ellas realmente nos envía mas información. Por ello utilizamos la ***tasa de información*** con la que podemos saber realmente quien nos da mas información, ya que si dos fuentes tienen la misma *esperanza de información* pero una de ellas tarda menos tiempo en enviar un símbolo, tendremos una mayor *tasa de información* para esta fuente y podremos decir que esta fuente tiene mejor tasa de información a la hora de comprar dos o mas fuentes.

   Para el calculo de la última parte del ejercicio necesitaremos la formula del calculo de la tasa:

   $$R(S) = \frac{H(S)}{\tau}$$

   donde τ = tiempo que se tarda en enviar un símbolo,

   $$H(S) = E\{I(S)\}$$

   y,

   $$I(S) = -\log_2(p(S))$$

   y,

   $$E\{I(S)\} = -\sum_{i=1}^{N} p(S_i) \cdot \log_2(p(S_i))$$

   Por lo que una vez conocemos estos datos, más la información del enunciado, tenemos que:

   - Fuente “A”:

     $$R(S) = \frac{0{,}8}{0{,}003} = 266{,}67\text{bps.}$$

   - Fuente “B”:

     $$R(S) = \frac{0{,}5}{0{,}001} = 500\text{bps.}$$

   Por lo que finalmente podemos decir que la fuente “B” tiene mayor tasa de información que la fuente “A”.

2. **Un teléfono es capaz de muestrear la voz humana a 8KHz. sabiendo que 1Hz = 1seg.$^{-1}$ (frecuencia=1/tiempo), ¿cuál es el periodo de muestreo de un teléfono? En audio, la máxima audiofrecuencia perceptible para el oído humano está en torno a los 20 kHz. ¿A qué velocidad se debería muestrear (en Hz) para asegurar que se puede reconstruir la señal íntegra en un receptor? ¿A qué intervalo de muestreo corresponde dicha velocidad en Hz?**

   Como un teléfono muestrea a 8KHz = 8000Hz, tenemos que:

$$8000 = \frac{1}{t}$$

$$t = \frac{1}{8000} = 0{,}000125\text{s} = 0{,}125\text{ms}$$

Por lo que podemos decir que un telefono muestra cada 0.125 milisegundos.

Sabemos que el Teorema de muestreo de Nyquist establece que:

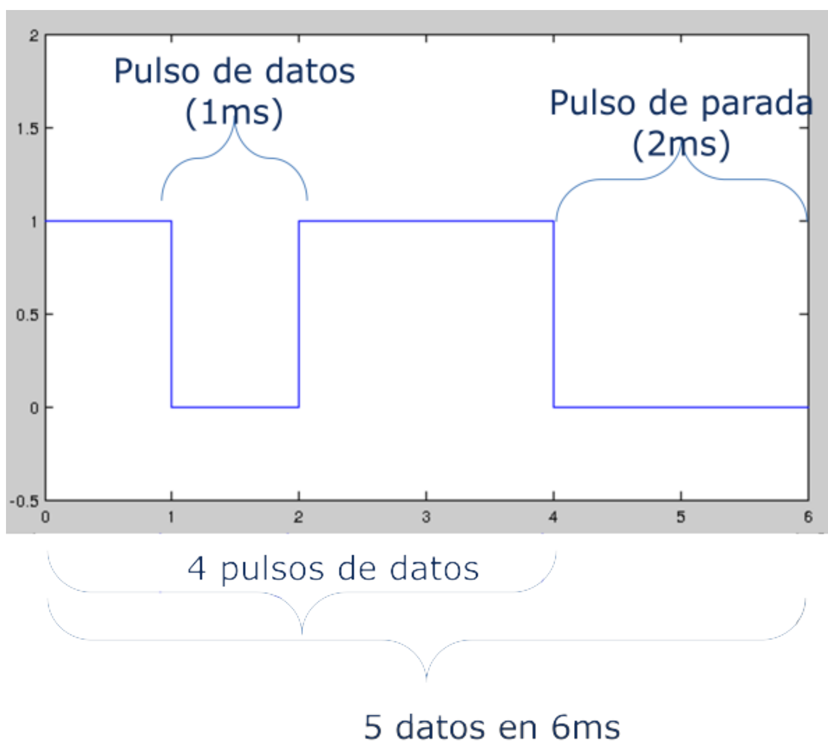
$$s \leq 2 \cdot B$$

Lo que nos dice que para una velocidad maxima $\boldsymbol{s}$, es necesario que el canal disponga de un ancho de banda superior o igual al doble de $\boldsymbol{s}$.

Esto implica que para medir datos que se envían cada τ instantes de tiempo, el receptor tiene que muestrear a como mucho $\frac{\tau}{2}$.

Tenemos que el oído humano esta en torno a los 20KHz = 20000Hz por lo que tenemos que $\tau = 1/20000 = 5\text{x}10^{-5}\text{s} = 0{,}05\text{ms}$ y esto nos dice que hay que muestrear cada $\frac{0{,}05}{2} = 0{,}025\text{ms} = 40\text{KHz}$

3. **Supongamos que tenemos el siguiente diagrama que modela una comunicación desde un emisor a un receptor:**



Responda a las siguientes cuestiones:

- **¿Cuál es el periodo de envío de una señal de datos?**

  Se envía una señal de datos cada 1ms.

- **¿Cuál es la velocidad de señalización del canal?**

Si consideramos sólo el canal y los bits de información:

$$r = \frac{1}{\tau}$$

por lo que,

$$r = \frac{1}{0{,}001} = 1000 \text{ baud.}$$

Pero si consideramos el mensaje íntegro, entonces:

$$r = \frac{n}{T}$$

donde $n =$ numero de datos de información de la trama y $T =$ al tiempo total de la trama. Por lo que,

$$r = \frac{5}{0{,}006} \approx 833 \text{ baund.}$$

- **¿Cuál es la velocidad máxima posible?**

$$s = \frac{1}{\tau_{min}}$$

por lo que,

$$s = \frac{1}{0{,}001} = 1000 \text{ baud.}$$

- **¿Cuál es la capacidad del canal?**

$$c = s \cdot \log_k(m)$$

donde $m =$ al numero de simbolos diferentes que haya y $k =$ al numero de niveles del canal. Por lo que,

$$c = 1000 \cdot \log_2(2) = 1000 \cdot 1 = 1000\text{bps}$$

- **Basándonos sólo en los datos que aporta la figura, ¿cuál es la información aportada por recibir una señal de datos "1"? ¿y la información de recibir una señal de datos "0"? ¿Cuál es la esperanza de información de la fuente?**

Basándonos solo en la imagen, podemos ver que tenemos 1 pulso de información en "0" y 3 pulsos en "1", por lo que asumimos que $p(\text{“}1'') = 0{,}75$ y $p(\text{“}0'') = 0{,}25$.

Por lo que tenemos que:

$$I(\text{“}1'') = -\log_2 0{,}75 = 0{,}415$$

$$I(\text{“}0'') = -\log_2 0{,}25 = 2$$

$$H(S) = E\{I(S)\} = \sum_{i=1}^{N} p(S_i) \cdot I(S_i) = 0{,}75 \cdot 0{,}415 + 0{,}25 \cdot 2 = 0{,}812$$

- **¿Cuál es la tasa de información del sistema?**

En un principio tenemos que:

$$R = \frac{H(S)}{\tau} = \frac{0{,}812}{0{,}001} = 812\text{bps}$$

Sin embargo, hemos transmitido 4 pulsos de información en 6ms, de los cuales los dos últimos no han aporta nada, por lo tanto:

$$R = k \cdot \frac{H(S)}{T}$$

donde k = numero de simbolos de informacion y T = tiempo total que se tarda en enviar la trama. Asique tenemos que:

$$R = 4 \cdot \frac{0{,}812}{0{,}006} \approx 541\text{bps}$$

- **¿Cuál debería ser el periodo de muestreo del receptor para poder reconstruir fielmente la señal en el destino?**

  $$\frac{\tau}{2} = \frac{1}{2} = 0{,}5\text{ms}$$

  Por lo que tenemos que muestrear como maximo a 0.5ms.

4. **Supongamos que tenemos un sistema de comunicaciones donde, en cada paso, un emisor envía una trama de 10 datos binarios por un canal con 2 niveles a un receptor en 14ms. De esos 10 datos, 6 son bits de información, y están organizados como sigue:**

   - **Bits 0 y 1: El primer símbolo enviado por el emisor.**
   - **Bit 2: Símbolo de parada (indica fin de envío del primer símbolo).**
   - **Bits 3 y 4: El segundo símbolo enviado por el emisor.**
   - **Bit 5: Símbolo de parada (indica fin de envío del primer símbolo).**
   - **Bits 6 y 7: El tercer símbolo enviado por el emisor.**
   - **Bit 8: Símbolo de parada (indica fin de envío del primer símbolo).**
   - **Bit 9: Símbolo de fin del mensaje (indica el fin de envío de la trama).**

   Cada bit de datos tarda en enviarse 1ms. Los bits de fin de símbolo y de fin de mensaje tardan en enviarse 2ms. Responda a las siguientes cuestiones:

   - **¿Cuál es el periodo de envío de una señal de datos?**

     En esta pregunta podemos responder de dos formas distintas, respondiendo sobre el canal o respondiendo sobre el emisor, sin embargo, lo haré sobre el canal, ya que es como lo hemos hecho hasta ahora.

     Por lo que como estamos hablando del canal, tenemos que cada señal de datos se envía cada 1 milisegundo.

   - **¿Cuál es la velocidad de señalización del canal?**

     Podemos ver que tardamos 1ms en enviar un bit de datos, por lo que tenemos que $\tau = 1\text{ms}$

     $$r_{teorica} = \frac{1}{0{,}001} = 1000\text{baud.}$$

     $$r_{real} = \frac{6}{0{,}014} \approx 428\text{baud.}$$

   - **¿Cuál es la velocidad máxima posible?**

     $$s = \frac{1}{0{,}001} = 1000\text{baud.}$$

   - **¿Cuál es la capacidad del canal?**

     $$c = 1000 \cdot \log_2(2) = 1000\text{bps}$$

- **Suponiendo que la probabilidad de enviar una señal de datos "0" y una señal de datos "1" es la misma, p("0")= p("1")= 0,5, ¿Qué aporta más información, enviar un 0 o enviar un 1? ¿Cuál es la esperanza de información de la fuente?**

$$p(\text{“}1'') = -\log_2 0{,}5 = 1$$

$$p(\text{“}0'') = -\log_2 0{,}5 = 1$$

$$H(S) = E\{I(S)\} = \sum_{i=1}^{2} p(S_i) \cdot I(S_i) = 0{,}5 * 1 + 0{,}5 * 1 = 1$$

  Todos los símbolos son equiprobables, existe la máxima incertidumbre posible.

- **¿Cuál es la tasa de información del sistema?**

$$R_{teorica} = \frac{1}{0{,}001} = 1000\text{bps}$$

$$R_{real} = 6 \cdot \frac{1}{0{,}014} \approx 428\text{bps}$$

- **¿Cuál debería ser el periodo de muestreo del receptor para poder reconstruir fielmente la señal en el destino?**

$$\frac{\tau}{2} = \frac{1}{2} = 0{,}5\text{ms}$$

  Por lo que tenemos que muestrear como maximo a 0,5ms.

5. **Suponga que el número de mujeres que vive más 80 años supera en una proporción 3/1 al número de hombres que viven por encima de esa edad (por cada hombre que supera los 80 años, hay 3 mujeres). ¿Cuánta cantidad de información se gana sabiendo que un hombre ha superado los 80 años?**

   Tenemos que por cada 3 mujeres tenemos un hombre, esto quiere decir que de 4 personas, la probabilidad de que una de ellas sea hombre es de 0,25 por lo que de que sea una mujer es 0,75.

   Denominamos "s1" al suceso de que una persona que pasa los 80 años sea mujer y "s2" al suceso de que sea un hombre,

$$I(s1) = -log_2(0{,}75) = 0{,}41$$

$$I(s2) = -log_2(0{,}25) = 2$$

   Lo que nos dice que la cantidad de informacion que nos ofrece la noticia de que un hombre es mayor de 80 años es mayor que la misma noticia pero de una mujer. Esto es lógico ya que como el suceso de ser un hombre que supera los 80 años es menos probable, en el caso de ocurrir, nos aporta más informacion que en el caso de la mujer.

6. **Indique los diferentes tipos de canales que conoce. Describa ejemplos de los mismos y en qué se diferencian unos de otros.**

   Los canales se dividen esencialmente en dos grupos mayoritarios que son los **caneles con ruido** y **canales sin ruido**.

   La diferencia esencial entre ellos es que en **un canal sin ruido** (o canal ideal) si una fuente genera k símbolos, el receptor siempre recibe exactamente los mismos símbolos, sin errores.

Tendríamos que $C = s{*}\log(m)$, lo que nos dice que la capacidad del canal está completamente aprobechada y siempre se utiliza en canal completo parar enviar información.

Por otra parte, **en canales con ruido**, no tenemos las caracteristicas anteriores y esto hace que cuando un emisor envíe k simbolos a un receptor, éste no tiene porque recibir exactamente los símbolos que realmente se enviaron.

Para tener diferenciados los distintos tipos de canales se crearon las siguientes subcategorías:

- **Canales deterministas**: Son canales donde la entrada determina unívocamente la salida, es decir, desde una fuente siempre se diferencia qué simbolo se está enviando, pero el receptor tiene un mismo símbolo para referenciar a varios del emisor, lo que hace que no pueda diferenciar exactamente de qúe simbolo proviene.
- **Canales sin pérdida**: Son canales que si conoces la salida, conoces unívocamente la entrada, es dedcir, el receptor sabe exactamente qué simbolo se le está enviando del emisor, aunque éste puede con un mismo simbolo enviado, que el receptor reciba símbolos distintos para el mismo simbolo.
- **Canales sin ruido**: Es simultáneamente determinista y sin pérdida, es decir, en estos canales tanto el receptor como el emisor saben exactamente que envían y reciben ya que la relación entre los símbolos de uno y otro son únicas.
- **Canales simétricos**: Todas las filas de la matriz de codificación contienen los mismos valores pero en orden distinto.
- **Canales inútiles**: La probabilidad de recibir un símbolo es la misma que de recibir otro, esto hace que el receptor no sepa distinguir qúe simbolo se le ha enviado y por lo tanto hace el canal inútil.

7. **Indique los diferentes tipos de códigos que conoce. Describa en qué se diferencian los códigos (unos de otros). Ponga ejemplos de cada uno de los códigos.**

   Podemos diferenciar varios tipos de códigos:

   - **Según el número de símbolos del alfabeto a utilizar**: Estos códigos pueden ser *Binarios, Ternarios...* dependiendo del número de símbolos. Por ejemplo un código *Binario* sería {0,1}, mientras que un tipo de código *Ternario* sería {1,2,3}
   - **Según la longitud que tenga una palabra transmitida**: Estos códigos a su vez se dividen en dos subtipos, que son códigos *Uniformes* y códigos *No uniformes*. La diferencia entre ellos es básicamente que un código *Uniforme* tiene todas las palabras de la misma longitud como por ejemplo el código ASCII, mientras que un código *No uniforme* tiene palabras con distinta longitud como puede ser el código Morse.
   - **Códigos de traducción única**: Estos códigos se caracterizan porque a la hora de decodificarlos, sólo tienen una traducción única y no puede haber inconsistencias. Por ejemplo, el código $m_1 = 0$, $m_2 = 01$ es un código de traducción única ya que cualquier mensaje compuesto por estos dos, sólo se puede decodificar de una única forma.
   - **Códigos instantáneos**: Estos códigos son aquellos en los que ninguna palabra comienza utilizando valores de otra palabra. Estos códigos son un subconjunto de los códigos de

traducción única. Por ejemplo $m_1 = 0$, $m_2 = 10$, $m_3 = 110$, $m_4 = 1110$ es un código de tracción única ya que ninguna palabra empieza utilizando ninguna otra.

- **Códigos óptimos**: Estos códigos son los códigos que como su nombre indica son óptimos para codificar los mensajes, es decir, se utilizan el mínimo número de símbolos para codificar todos los mensajes. Por ejemplo, asumiento que nuestro lenguaje tiene únicamente 2 símbolos, un código óptimo sería $m_1 = 0$, $m_2 = 1$, mientras que $m_1 = 01$, $m_2 = 10$ no lo sería ya que desperdiciamos valores.

8. **Se dispone de un canal que puede codificar 3 niveles por instante de tiempo. Diseñe un código óptimo que permita enviar 9 símbolos (alfabeto de la fuente) por el canal, asumiendo que todos los símbolos son equiprobables. Demuestre que el código es óptimo.**

   Para ello, como sabemos que tenemos 3 niveles por instante de tiempo, en vez de utilizar bits para codificar los mensajes, utilizaremos trips. Por lo tanto vamos a calcular el número de trips que necesitamos para codificar los 9 mensajes. Para ello, utilizaremos el logaritmo.

   $$\log_3(9) = 2 \text{ trips.}$$

   Por lo que tenemos que para codificar 9 símbolos con 3 niveles necesitamos 2 trips.

   Ahora, tenemos entonces que nuestro alfabeto será {*1,2,3*}, asíque tendríamos:

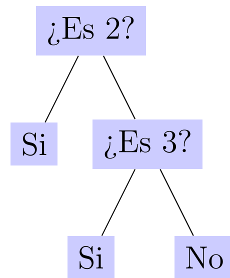
   - $s0 = 11$
   - $s1 = 12$
   - $s2 = 13$
   - $s3 = 21$
   - $s4 = 22$
   - $s5 = 23$
   - $s6 = 31$
   - $s7 = 32$
   - $s8 = 33$

   Podemos ver como utilizamos todas las diferentes combinaciones para codificar todos los símbolos y no nos queda ninguna suelta, por ello, podemos comprobar que este código es óptimo.

9. **Un juego entre dos amigos consiste en que suponen que la cara de la moneda vale 1 y la cruz vale 2. Uno de ellos lanza una moneda trucada (p(Cara)= 0.8, p(Cruz)= 0.2) al aire 2 veces, y suma el resultado de ambas tiradas. El otro amigo debe adivinar qué número ha calculado el primero. Desarrolle una estrategia que permita realizar el mínimo número de preguntas de respuesta "Sí/No" para averiguar la respuesta.**

   En este caso tendríamos 3 tipo de sucesos, que salga un 2 (dos caras) un 3 (una cara y una cruz) o un 4 (dos cruces), claramente lo más probable es que salga un 2 ya que sería que saliese dos veces cara, cosa que es bastante probable, la segunda opción seria un 3 que tiene

1 cara que es muy probable y una cruz. Por último el caso menos probable que sería un 4 sacando dos cruces. Por lo que mi árbol de decisiones sería:



Como mucho, en el peor caso haremos 2 preguntas, pero como lo más probable es que sea 2, lo más seguro es que sólo tengamos que realizar una pregunta.