Curso:
EL271 Codificación de Fuente y de Canal
Unidad 1: TEORÍA DE LA INFORMACIÓN Y CODIFICACIÓN DE FUENTE
Semana 3, Sesión 3

CARLOS VALDEZ VELÁSQUEZ-LÓPEZ, DR. ING

2022-02

Logro de la sesión

Al finalizar la sesión, el alumno explica los conceptos más importantes sobre la transmisión de información sobre canales discretos, la información mutua, la capacidad del canal discreto y continuo, y el Teorema de Shannon de la capacidad del canal

Contenido

1) Transmisión de información en canales discretos, información mutua
2) El canal binario simétrico (BSC)
3) La capacidad del canal discreto
4) El Teorema de Shannon de la capacidad del canal
5) La capacidad del canal continuo

1) Transmisión de información en canales discretos, información mutua

- Considere el sistema de transmisión de información discreto:

- Una fuente discreta selecciona símbolos de un alfabeto X para transmitirlos por el canal
- Idealmente, el canal debería reproducir en el destino los mensajes emitidos por la fuente
- El ruido y otras perturbaciones alteran los símbolos transmitidos resultando un alfabeto Y de símbolos diferente
- Se desea medir la información transmitida en este caso

1) Transmisión de información en canales discretos, información mutua

- Para tratar con ambos alfabetos Xe Y, se define la notación:

$P(x_i)$ es la probabilidad de que la fuente seleccione el símbolo x_i para ser transmitido

$P(y_i)$ es la probabilidad de que el símbolo y_i sea recibido en el destino

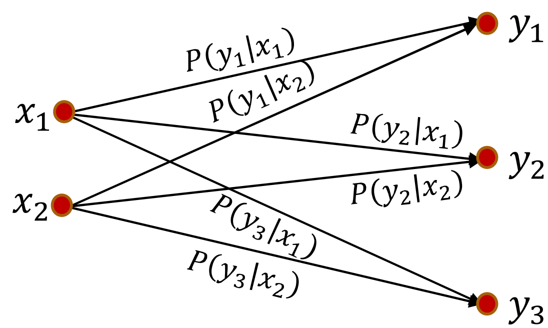
$P(x_i y_j)$ es la probabilidad conjunta de que se transmita el símbolo x_i y que se reciba el símbolo y_j

$P(x_i|y_j)$ es la probabilidad condicional de que x_i fue transmitido, dado que y_j fue recibido

$P(y_j|x_i)$ es la probabilidad condicional de que y_j es recibido, dado que x_i fue transmitido (se conoce también como la probabilidad de transición del canal ("forward transition probability")

- Se asume que el canal es invariante en el tiempo y sin memoria, de modo que las probabilidades condicionales son independientes del tiempo y de los símbolos previamente transmitidos

Ejemplo: 2 símbolos de la fuente y 3 símbolos en el destino, en canal con ruido



- El sistema pretende entregar $y_j = y_1$ cuando se transmite $x_i = x_1$, y entregar $y_j = y_2$ cuando se transmite $x_i = x_2$
- Por lo tanto, la probabilidad de error de símbolo está dada por $P\left(y_j \middle| x_i\right)$, para $j \neq i$
- Ejemplo: supongamos que $x_1 = 10$ y $x_2 = 15$, y que el ruido puede tomar 3 valores, $n_1 = 5$, $n_2 = 3$ y $n_3 = 2$, entonces si se transmite x_1, se recibiría $y_1 = 10 + 5 = 15$, $y_2 = 10 + 3 = 13$, $y_3 = 10 + 2 = 12$

La información mutua

- La cantidad de información transferida cuando se transmite x_i y se recibe y_j está dada por:

$$I(x_i; y_j) \triangleq \log \frac{P(x_i/y_j)}{P(x_i)} \text{ bits} \quad (1)$$

- Supongamos que se tiene un canal ideal libre de ruido, tal que cada y_j identifica únicamente a un x_i particular, entonces $P\left(x_i \middle| y_j\right) = 1$, tal que:

$$I\left(x_i; y_j\right) = \log 1/P(x_i) \quad (2)$$

En tal caso, la transferencia de información es igual a la auto información de x_i

- Por otro lado, suponga que el ruido del canal tiene un efecto tal que y_j no guarda relación con x_i, entonces $P\left(x_i \middle| y_j\right) = P(x_i)$, de modo que:

$$I\left(x_i; y_j\right) = \log 1 = 0 \quad (3)$$

en cuyo caso no se transmite ninguna información

- Los casos anteriores son extremos e intuitivamente y respaldan la definición de que la información mutua es una medida de la transferencia de información

La información mutua

- La mayoría de los canales de transmisión cae en algún lugar entre los extremos de perfecta transferencia y transferencia nula. Para analizar el caso general, se define la información mutua promedio:

$$I(X;Y) \triangleq \sum_{x,y} P\left(x_i\, y_j\right) I\left(x_i; y_j\right) \text{ (4)}$$

$$= \sum_{x,y} P\left(x_i\, y_j\right) \log \frac{P(x_i/y_j)}{P(x_i)} \text{ bits/símbolo (5)}$$

en donde los índices de la sumatoria indican que el promedio estadístico se toma sobre ambos alfabetos

- La cantidad $I(X;Y)$ representa la cantidad de información promedio de la fuente ganada por símbolo recibido, a diferencia de la información promedio por símbolo de la fuente representado por la entropía de la fuente $H(X)$

La información mutua

Lectura opcional

- Por otro lado, teniendo en cuenta las siguientes expresiones:

$$P(x_i\, y_j) = P(x_i|y_j)P(y_j) = P(y_j|x_i)P(x_i) \quad (6)$$

$$P(x_i) = \sum_y P(x_i\, y_j) \quad P(y_j) = \sum_x P(x_i\, y_j) \quad (7)$$

- Las reemplazamos en $I(X;Y)$:

$$I(X;Y) = \underbrace{\sum_{x,y} P(x_i\, y_j)\log\frac{1}{P(x_i)}}_{(1)} - \underbrace{\sum_{x,y} P(x_i\, y_j)\log\frac{1}{P(x_i|y_j)}}_{(2)} \quad (8)$$

- El primer término (1):

$$\sum_x\left[\sum_y P(x_i\, y_j)\right]\log\frac{1}{P(x_i)} = \sum_x P(x_i)\log\frac{1}{P(x_i)} = H(X) \quad (9)$$

- El segundo término (2) se define como la equivocación:

$$H(X|Y) \triangleq \sum_{x,y} P(x_i\, y_j)\log\frac{1}{P(x_i|y_j)} \quad (10)$$

La información mutua

- Las reemplazamos en $I(X;Y)$: $I(X;Y) = H(X) - H(X|Y)$ (11)
- Esta expresión se interpreta como la transferencia de información promedio por símbolo, y es igual a la entropía de la fuente menos la equivocación
- La equivocación representa la pérdida de información en un canal con ruido
- Otra perspectiva de la transferencia se obtiene a partir de la expresión (6), tal que:

$$P(x_i|y_j)/P(x_i) = P(y_j|x_i)/P(y_j) \quad (12)$$

➢ En consecuencia: $I(X;Y) = I(Y;X)$, e intercambiando X y Y se obtiene:

$$I(Y;X) = H(Y) - H(Y|X) \quad (13)$$

Lectura opcional

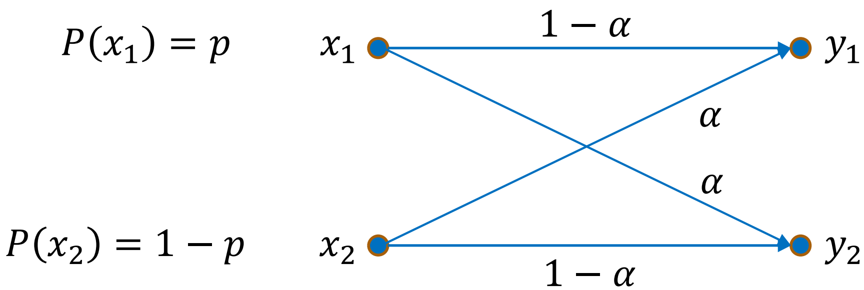
➢ En donde: $H(Y) = \sum_y P(y_j) \log \frac{1}{P(y_j)}$ (14) y $H(Y|X) \triangleq \sum_{x,y} P(x_i\, y_j) \log \frac{1}{P(y_j|x_i)}$ (15)

➢ La expresión (13) se interpreta como la entropía del destino $H(Y)$ menos la entropía del ruido $H(Y|X)$ añadida por el canal (lo cual se justifica porque las probabilidades de transición $P(y_j|x_i)$ incluyen las probabilidades de error de símbolo

2) El canal binario simétrico (Binary Symmetric Channel, BSC)

Lectura opcional

- El modelo del BSC se muestra a continuación:



- Se tiene 2 símbolos de la fuente con probabilidades $P(x_1) = p$ y $P(x_2) = 1 - p$ (16)
- Asimismo, se tiene 2 símbolos de destino, con probabilidades de transición:

$$P(y_1|x_2) = P(y_2|x_1) = \alpha \quad (17)$$

$$P(y_1|x_1) = P(y_2|x_2) = 1 - \alpha \quad (18)$$

2) El canal binario simétrico (Binary Symmetric Channel, BSC)

- El modelo anterior podría representar cualquier sistema de transmisión en el que los errores son estadísticamente independientes y donde las probabilidades de error α son iguales para ambos símbolos, tal que la probabilidad de error promedio por símbolo es:

$$P_e = P(x_1)P(y_2|x_1)+P(x_2)P(y_1|x_2) = p\alpha + (1-p)\alpha = \alpha \quad (19)$$

- Ya que se conocen las probabilidades de transición, utilizaremos $I(Y;X) = H(Y) - H(Y|X)$ para calcular la información mutua en términos de p y α

- $H(Y)$ se obtiene considerando la salida del canal como una fuente binaria con probabilidades de símbolo $P(y_1)$ y $P(y_2) = 1 - P(y_1)$, en cuyo caso:

$$H(Y) = \Omega[P(y_1)] \quad (20)$$

- En donde $\Omega(\ )$ es la función de entropía binaria y:

$$P(y_1) = P(y_1|x_1)P(x_1)+P(y_1|x_2)P(x_2) = (1-\alpha)p + \alpha(1-p) = \alpha + p - 2\alpha p \quad (21)$$

2) El canal binario simétrico (Binary Symmetric Channel, BSC)

Lectura opcional

- En cuanto a $H(Y|X)$, se reemplaza $P(x_i\, y_j) = P(y_j|x_i)P(x_i)$ en (15) tal que:

$$H(Y|X) = \sum_{x,y} P(y_j|x_i)P(x_i)\log\frac{1}{P(y_j|x_i)} = \sum_{x} P(x_i)\sum_{y} P(y_j|x_i)\log\frac{1}{P(y_j|x_i)}$$

$$= P(x_1)\sum_{y} P(y_j|x_1)\log\frac{1}{P(y_j|x_1)} + P(x_2)\sum_{y} P(y_j|x_2)\log\frac{1}{P(y_j|x_2)}$$

$$= P(x_1)\left[P(y_1|x_1)\log\frac{1}{P(y_1|x_1)} + P(y_2|x_1)\log\frac{1}{P(y_2|x_1)}\right] +$$

$$P(x_2)\left[P(y_1|x_2)\log\frac{1}{P(y_1|x_2)} + P(y_2|x_2)\log\frac{1}{P(y_2|x_2)}\right]$$

$$= p\left[(1-\alpha)\log\frac{1}{(1-\alpha)} + \alpha\log\frac{1}{\alpha}\right] + (1-p)\left[(1-\alpha)\log\frac{1}{(1-\alpha)} + \alpha\log\frac{1}{\alpha}\right]$$

$$= (1-\alpha)\log\frac{1}{(1-\alpha)} + \alpha\log\frac{1}{\alpha} = \Omega(\alpha) \quad (22)$$

- Entonces: $I(Y;X) = H(Y) - H(Y|X) = \Omega(\alpha + p - 2\alpha p) - \Omega(\alpha)$ (23)

2) El canal binario simétrico (Binary Symmetric Channel, BSC)

Lectura opcional

- Lo cual significa que la transferencia de información depende de la probabilidad de error α y **de la probabilidad de la fuente $\boldsymbol{p}$**

- Por lo tanto, si el ruido es pequeño, $\alpha \ll 1$ y en consecuencia:

$$I(X;Y) \approx \Omega(p) = H(X) \quad (24)$$

- Por el contrario, si el ruido es muy grande, $\alpha = 1/2$, y en ese caso:

$$I(X;Y) = \Omega(\alpha + p - 2\alpha p) - \Omega(\alpha) = \Omega\left(\tfrac{1}{2} + p - 2 \cdot \tfrac{1}{2} \cdot p\right) - \Omega\left(\tfrac{1}{2}\right) = 0 \quad (25)$$

3) La capacidad del canal discreto

- Un canal de comunicación tiene alfabetos de la fuente (como x_1 y x_2) y de destino (como y_1, y_2 y y_3) fijos, así como probabilidades de transición fijas (como $P\left(y_j \middle| x_i\right)\ i = 1,2\ \ j = 1,2,3$)
- Por lo tanto, las únicas cantidades variables en $I(X;Y)$ son las probabilidades de fuente $P(x_i)$
- En tal sentido, la máxima transferencia de información requiere la estadística específica de la fuente, obtenidas a través de la codificación de fuente
- Sea C_s la máxima cantidad de información transferida por símbolo, denominada la capacidad del canal:

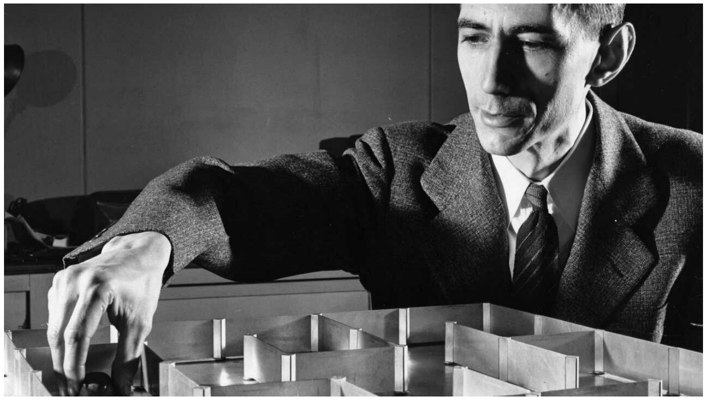
$$C_s \triangleq \underbrace{\max}_{P(x_i)} I(X;Y) \quad \text{bits/símbolo}$$

- La capacidad del canal del también se mide en términos de la velocidad de información. Si s representa la máxima velocidad de símbolos permitida por el canal, entonces:

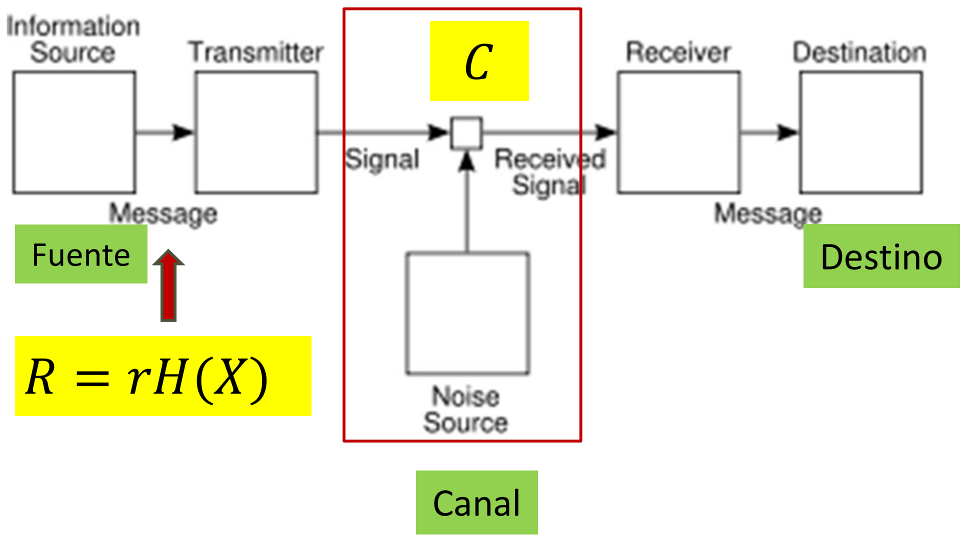
$$C = sC_s \quad \text{bits/segundo}$$

4) El Teorema de Shannon de la capacidad del canal

"Si un canal tiene capacidad C y una fuente tiene velocidad de información $R \leq C$, entonces existe un sistema de codificación tal que la salida de la fuente puede transmitirse a través del canal con una probabilidad de error P_e arbitrariamente pequeña. Por el contrario, si $R > C$, no es posible transmitir la información sin errores"



4) El Teorema de Shannon de la capacidad del canal



Fuente: “A Mathematical Theory of Communications”, C.E. Shannon, The Bell System Technical Journal, October 1948

5) La capacidad del canal continuo

- Para un canal continuo ideal limitado en banda a B Hz, con relación señal a ruido S/N, la capacidad de canal C:

$$C = B \ \log(1 + S/N) \ \text{bits/segundo}$$

- Expresión compacta que define la capacidad del canal en función de dos parámetros
- S/N y B, son limitaciones físicas inherentes al canal
- S/N: relación señal a ruido disponible
- B: ancho de banda en Hz

Formas de alcanzar la capacidad del canal

- Dicho teorema es no constructivo pues no indica cómo deben diseñarse los códigos
- Sin embargo, existe una deducción matemática sobre la probabilidad de error, que demuestra que $P_e \to 0$
- Mediante señales ortogonales, por ejemplo, M-FSK: tienen la gran desventaja del ancho de banda que ocupan
- Mediante técnicas de codificación de canal, de acuerdo a lo indicado por C.E. Shannon en su Teorema

Conclusiones