Canais Discretos Sem Memória

Teoria da Informação - AULA 15
Profª. Verusca Severo

Universidade de Pernambuco
Escola Politécnica de Pernambuco

27 de agosto de 2021

Introdução

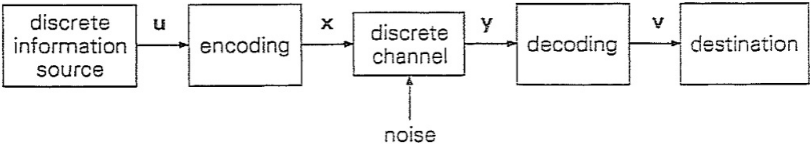
- Na primeira unidade de nosso curso, nos concentramos na definição da quantidade mínima de unidades de informação binária (*bit*) por símbolo necessário para representar completamente uma fonte de informação (a eficiência da representação da informação gerada pela fonte).

- Agora, a pergunta que nos surge é: "Qual é o valor máximo da taxa de transmissão que garante confiabilidade da comunicação através de um canal ruidoso?".

- Na aula de hoje, iremos abordar a essa questão usando o modelo de canal discreto sem memória.

Canal Discreto

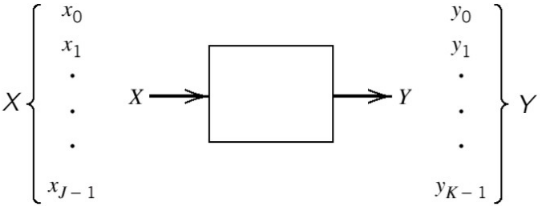
- Um canal discreto é um modelo estatístico que a um alfabeto de entrada representado por uma variável aleatória discreta X faz corresponder, de acordo com uma lei de transição, uma variável aleatória discreta Y.



- O sistema é discreto pois X e Y contém um número finito de palavras-código.

Canal Discreto

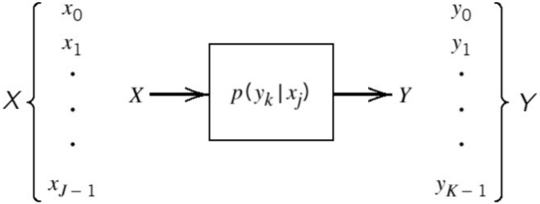
- A variável aleatória discreta X modela o alfabeto de entrada do canal, e Y modela o alfabeto de saída.



- Obs.: Em geral, a cardinalidade do alfabeto de saída pode ser diferente da do alfabeto de entrada.

Canal Discreto Sem Memória

- um canal discreto **sem memória** é determinado pelo mecanismo estatístico que descreve o transporte de informação entre a fonte e o destinatário e que se define pelo conjunto de probabilidades condicionais.



- O sistema é sem memória pois a saída não depende da saída anterior.

Canal Discreto Sem Memória

- O sinal de entrada é modelado por uma variável aleatória:

$$X = \{x_0, x_1, \ldots, x_{J-1}\}$$

- O sinal de saída, que é uma versão ruidosa do sinal de entrada, é modelado por uma variável aleatória:

$$Y = \{y_0, y_1, \ldots, y_{K-1}\}$$

- Para caracterizar o canal é necessário conhecer as probabilidades de transição $p(y_k|x_j)$.

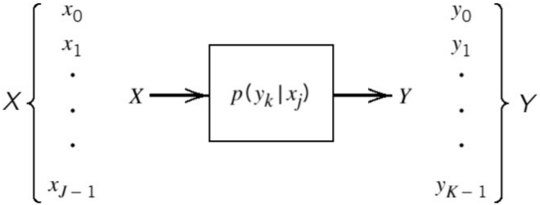
Canal Discreto Sem Memória

- Temos que:

$$p(y_k|x_j), \quad k = 0, 1, \dots, K-1; \quad j = 0, 1, \dots, J-1$$

com

$$0 \leq p(y_k|x_j) \leq 1$$



- Note que cada símbolo de saída só depende de um símbolo de entrada e não de uma sequência. Daí a designação de canal sem memória.

Canal Discreto Sem Memória

- Uma forma de caracterizar um canal discreto sem memória é através da matriz de transição (ou matriz de canal), que apresenta as várias probabilidades de transição do canal:

$$P = \begin{bmatrix} p(y_0|x_0) & p(y_1|x_0) & \dots & p(y_{K-1}|x_0) \\ p(y_0|x_1) & p(y_1|x_1) & \dots & p(y_{K-1}|x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p(y_0|x_{J-1}) & p(y_1|x_{J-1}) & \dots & p(y_{K-1}|x_{J-1}) \end{bmatrix}$$

  - Cada linha de P corresponde a uma entrada de canal fixa, $x_j, j = 0, 1, \dots, J-1$.
  - Cada coluna de P corresponde a uma saída de canal fixa, $y_k, k = 0, 1, \dots, K-1$.

Canal Discreto Sem Memória

- Se um símbolo em particular x_j for inserido na entrada do canal, haverá um símbolo na saída com uma dada probabilidade, ou seja, a emissão de x_m pode corresponder a recepção de qualquer um dos símbolos do alfabeto de saída.

- Em geral:

$$\forall j = 0, 1, \ldots, J-1 : p(y_k|x_j) \neq 0, k \neq j$$

  ou seja, o processo de transmissão está sujeito a erros.

- A ocorrência de erros de transmissão decorre do fato de o canal ser ruidoso.

Canal Discreto Sem Memória

- Assim, à emissão do símbolo particular x_j pode corresponder a recepção de qualquer um dos símbolos do alfabeto de saída. Logo:

$$\sum_{k=0}^{K-1} p(y_k|x_j) = 1 \quad \forall j$$

- Isso quer dizer que a soma dos elementos ao longo de qualquer linha da matriz é sempre igual a um.

Canal Discreto Sem Memória

- A medida de fiabilidade da transmissão através de um canal discreto é dada pela **probabilidade média de erro por símbolo**.

- Seja $p(x_j)$, com $j = 0, 1, \ldots, J-1$, a probabilidade de o símbolo x_j ser transmitido. Naturalmente, ocorre um erro de transmissão se o símbolo recebido for um qualquer y_k, com $k \neq j$, isto é,

$$P_e = \sum_{\substack{k=0 \\ k \neq j}}^{K-1} p(y_k) \tag{1}$$

Canal Discreto Sem Memória

- Sabemos que a distribuição de probabilidade conjunta das variáveis X e Y é dada por
$$p(x_j, y_k) = p(x_j).p(y_k|x_j)$$
- A partir da lei das probabilidades marginais pode-se obter a probabilidade de um dado símbolo y_k aparecer na saída do canal:
$$p(y_k) = \sum_{j=0}^{J-1} p(x_j, y_k) = \sum_{j=0}^{J-1} p(x_j).p(y_k|x_j) \tag{2}$$
- Isso quer dizer que, conhecendo a distribuição de probabilidade de X e a matriz de transição, podemos calcular as probabilidades dos vários símbolos de saída $p(y_k)$.

Canal Discreto Sem Memória

- Substituindo a equação (2) na equação (1), obtém-se:

$$P_e = \sum_{\substack{k=0 \\ k \neq j}}^{K-1} p(y_k)$$

$$P_e = \sum_{\substack{k=0 \\ k \neq j}}^{K-1} \sum_{j=0}^{J-1} p(x_j).p(y_k|x_j)$$

- Logo, a **probabilidade média de erro de transmissão por símbolo** é completamente determinada pelas probabilidades de transição.

Canal Discreto Sem Memória

- Uma maneira simples e eficiente de modelar o comportamento de um canal de comunicação binário é a partir de uma canal binário simétrico sem memória.

- O canal binário simétrico é um caso especial do canal discreto sem memória com $J = K = 2$, ou seja, o canal tem dois símbolos de entrada $(x_0 = 0, x_1 = 1)$ e dois símbolos de saída $(y_0 = 0, y_1 = 1)$

- Esse canal é dito simétrico porque a probabilidade de receber 1 supondo ter sido transmitido 0 é igual à probabilidade de receber 0 supondo ter sido transmitido 1, ou seja:

$$p(y_1 = 1|x_0 = 0) = p(y_0 = 0|x_1 = 1) = p$$

Assim, $P_e = p$ independente da distribuição de probabilidades.

- Se

$$p(y_1 = 1|x_0 = 0) = p(y_0 = 0|x_1 = 1) = p$$

  então

$$p(y_0 = 0|x_0 = 0) = 1 - p(y_1 = 1|x_0 = 0) = 1 - p$$

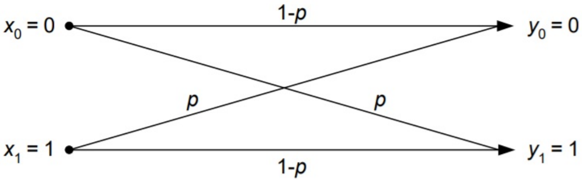
$$p(y_1 = 1|x_1 = 1) = 1 - p(y_0 = 0|x_1 = 1) = 1 - p$$

- Temos, então, a seguinte matriz de canal:

$$P = \begin{bmatrix} p(y_0|x_0) & p(y_1|x_0) \\ p(y_0|x_1) & p(y_1|x_1) \end{bmatrix} \Rightarrow P = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

Canal Discreto Sem Memória

- Diagrama de probabilidade de transição de um canal binário simétrico:



$$p(y_1 = 1|x_0 = 0) = p(y_0 = 0|x_1 = 1) = p$$

$$p(y_0 = 0|x_0 = 0) = p(y_1 = 1|x_1 = 1) = 1 - p$$

Canal Discreto Sem Memória

- Conceitos de Entropia:
  - $H(X)$ é uma medida de incerteza a respeito da entrada do canal antes de observarmos a saída (Mede a incerteza inicial associada à transmissão do símbolo).

  - $H(X|Y)$ é uma medida de incerteza a respeito da entrada do canal depois de observarmos a saída (Mede a incerteza final sobre a transmissão de X após ter sido recebido Y).
    - O canal,contento ruído é que cria incerteza em X dado que se conhece Y, alterando $H(X|Y)$.

  - $I(X;Y)$ é uma medida de incerteza a respeito da entrada do canal que é resolvida observando-se a saída do canal. Ou seja, determina a quantidade de informação obtida sobre uma ponta do canal, dado que foi observada a outra ponta do canal (é a quantidade de incerteza que é resolvível pelo canal).

$$I(X;Y) = H(X) - H(X/Y)$$

Canal Discreto Sem Memória

- Temos que:

$$I(X;Y) = H(X) - H(X/Y) \tag{3}$$

- Sabemos que:

$$H(X) = \sum_{j=0}^{J-1} p(x_j) \log_2 \left[\frac{1}{p(x_j)} \right] \tag{4}$$

e

$$H(X|Y) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j|y_k) p(y_k) \log_2 \left[\frac{1}{p(x_j|y_k)} \right] \tag{5}$$

Canal Discreto Sem Memória

- Substituindo as equações (4) e (5) na equação (3), temos:

$$I(X;Y) = H(X) - H(X/Y)$$

$$I(X;Y) = \sum_{j=0}^{J-1} p(x_j)\log_2\left[\frac{1}{p(x_j)}\right] - \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j|y_k)p(y_k)\log_2\left[\frac{1}{p(x_j|y_k)}\right]$$

- Sabemos que:

$$p(x_j|y_k)p(y_k) = p(x_j, y_k)$$

e que, pela probabilidade marginal:

$$p(x_j) = \sum_{k=0}^{K-1} p(x_j, y_k)$$

Canal Discreto Sem Memória

- Logo:

$$I(X;Y) = H(X) - H(X/Y)$$

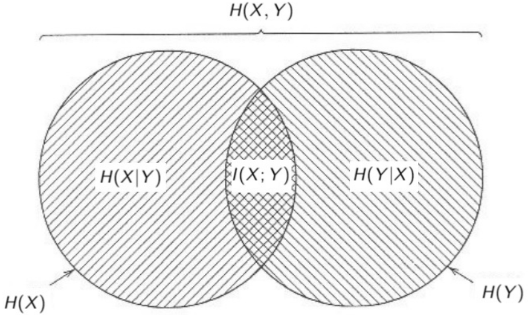
$$\text{I(X;Y)} = \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k)\log_2\left[\frac{1}{p(x_j)}\right] - \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k)\log_2\left[\frac{1}{p(x_j|y_k)}\right]$$

$$I(X;Y) = \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k)\left(\log_2\left[\frac{1}{p(x_j)}\right] - \log_2\left[\frac{1}{p(x_j|y_k)}\right]\right)$$

$$I(X;Y) = \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k)\log_2\left[\frac{\frac{1}{p(x_j)}}{\frac{1}{p(x_j|y_k)}}\right]$$

$$I(X;Y) = \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k)\log_2\left[\frac{p(x_j|y_k)}{p(x_j)}\right] \tag{6}$$

- O diagrama de Venn, com as todas as relações existentes de entropia e informação mútua, demonstra as relações entre as diversas quantidades introduzidas e que caracterizam a fonte, o canal, e a respectiva saída.



Canal Discreto Sem Memória

- Um canal discreto sem memória é caracterizado pelas probabilidades de transição $p(x_j|y_k)$.

- A informação mútua depende da distribuição de probabilidade de entrada (fonte) e da probabilidade de transição do canal.

$$I(X;Y) = \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \left[\frac{p(x_j|y_k)}{p(x_j)}\right]$$

$$I(X;Y) = \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j).p(y_k|x_j) \log_2 \left[\frac{p(x_j|y_k)}{p(x_j)}\right]$$

- Portanto, a informação mútua de uma canal depende não somente do canal, mas também da maneira como ele é usado.

Canal Discreto Sem Memória

- Então, temos:
$$I(X;Y) = H(X) - H(X/Y),$$
em que:
  - $H(X)$ é a **ENTROPIA DA FONTE**: é a incerteza média sobre a fonte de informação X.
  - $H(X/Y)$ é a **EQUIVOCAÇÃO**: é a incerteza média sobre a fonte depois de se observar a saída do canal, Y.
- Logo, a informação mútua é uma medida da informação que passou através do canal.
  - Quanto maior for a EQUIVOCAÇÃO menor é a quantidade de informação que passa através do canal.

Canal Discreto Sem Memória

- Para calcularmos $I(X;Y)$ precisamos conhecer a distribuição de probabilidade dos símbolos de entrada $p\{(x_j)\}$.

- A informação mútua portanto depende não apenas das probabilidades de transição do canal, $p\{(x_j/y_k)\}$, mas também dos valores $p\{(x_j)\}$, isto é, depende da maneira como o canal é utilizado.

- Sabemos que $H(X/Y) \leq H(X)$ (vimos isso na primeira unidade), isso quer dizer que a informação mútua é não-negativa, ou seja,

$$I(X;Y) \geq 0$$

  - O valor mínimo $I(X;Y) = 0$ ocorre se e somente se as X e Y forem estatisticamente independentes, i.e., $p\{(x_j/y_k)\} = p\{(x_j)\}\ \forall (j,k)$.

Canal Discreto Sem Memória

- Como a distribuição de probabilidade de entrada $p\{(x_j)\}$ independe do canal, podemos então maximizar a informação mútua $I(X;Y)$ do canal em relação a $p\{(x_j)\}$.

- Diante do exposto, define-se **Capacidade do canal** como o valor máximo da informação mútua $I(X;Y)$, obtido quando são variadas as probabilidades de entrada $p(x_j)$, com $j = 0, \ldots, J-1$, ou seja

$$C = \max_{\{p(x_j)\}} I(X;Y)$$

  - A capacidade C é medida em *bits* por utilização de canal ou *bits* por transmissão.

Canal Discreto Sem Memória

- A capacidade de canal depende apenas das probabilidades de transição que definem o canal.

- Otimizar C significa calcular para cada uma das J variáveis $p(x_j)$ que leva $I(X;Y)$ ao valor máximo.

$$C = \max_{\{p(x_j)\}} I(X;Y) = \max_{\{p(x_j)\}} \{H(X) - H(X|Y)\}$$

- A capacidade do canal somente será atendida se a fonte de entrada apresentar a distribuição adequada.
  - Para um canal ideal, temos:

$$H(X|Y) = 0 \quad \Rightarrow \quad C = \max_{\{p(x_j)\}} I(X;Y) = \max_{\{p(x_j)\}} \{H(X)\}$$

Canal Discreto Sem Memória

- Vejamos a característica de alguns tipos de canais:
  - Canal sem perdas:

$$H(X|Y) = 0 \quad \Rightarrow \quad I(X;Y) = H(X)$$

  - Canal inútil:

$$H(X|Y) = H(X) \quad \Rightarrow \quad I(X;Y) = 0$$

  - Canal simétrico:

$$I(X;Y) = ?$$