

# Teoría de la Información

- Clase práctica -

## Teoría de las Comunicaciones



DEPARTAMENTO  
DE COMPUTACION

Departamento de Computación  
Facultad de Ciencias Exactas y Naturales  
Universidad de Buenos Aires

23.08.2017

# Agenda

- 1 Introducción
- 2 Fuentes de información
- 3 Códigos
  - Generalidades
  - Tipos de códigos
  - Códigos sobre fuentes de memoria nula
- 4 Capacidad de canal
- 5 Delay

# Agenda

## 1 Introducción

## 2 Fuentes de información

## 3 Códigos

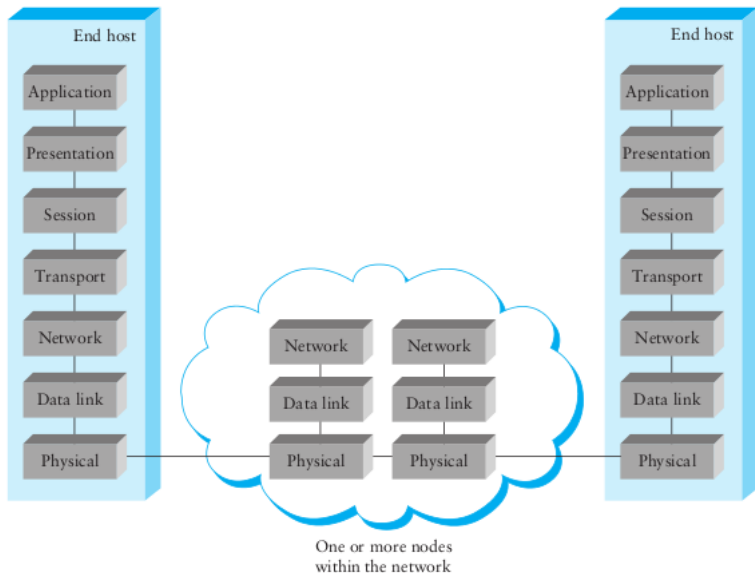
- Generalidades
- Tipos de códigos
- Códigos sobre fuentes de memoria nula

## 4 Capacidad de canal

## 5 Delay

# Miscelánea de temas organizativos de la materia

# Miscelánea de temas organizativos de la materia



# Agenda

## 1 Introducción

## 2 Fuentes de información

## 3 Códigos

- Generalidades
- Tipos de códigos
- Códigos sobre fuentes de memoria nula

## 4 Capacidad de canal

## 5 Delay

# Información y entropía

Información de un evento  $e$

$$I(e) = -\log_2 P(e) \text{ bits}$$

# Información y entropía

## Información de un evento $e$

$$I(e) = -\log_2 P(e) \text{ bits}$$

## Bit

Cantidad de información obtenida al especificar una de dos posibles alternativas igualmente probables



# Información y entropía

## Información de un evento $e$

$$I(e) = -\log_2 P(e) \text{ bits}$$

## Bit

Cantidad de información obtenida al especificar una de dos posibles alternativas igualmente probables

## Fuente de memoria nula

Fuente en la que cada emisión es estadísticamente independiente

# Información y entropía

## Información de un evento $e$

$$I(e) = -\log_2 P(e) \text{ bits}$$

## Entropía de una fuente $S$

$$H(S) = -\sum_{s \in S} P(s) \cdot \log_2 P_S(s)$$

## Bit

Cantidad de información obtenida al especificar una de dos posibles alternativas igualmente probables

## Fuente de memoria nula

Fuente en la que cada emisión es estadísticamente independiente

# Información y entropía

## Información de un evento $e$

$$I(e) = -\log_2 P(e) \text{ bits}$$

## Entropía de una fuente $S$

$$H(S) = -\sum_{s \in S} P(s) \cdot \log_2 P_S(s)$$

## Entropía bajo equiprobabilidad

$$H(S) = \log_2 |S|$$

## Bit

Cantidad de información obtenida al especificar una de dos posibles alternativas igualmente probables

## Fuente de memoria nula

Fuente en la que cada emisión es estadísticamente independiente

# Agenda

1 Introducción

2 Fuentes de información

3 **Códigos**

- Generalidades
- Tipos de códigos
- Códigos sobre fuentes de memoria nula

4 Capacidad de canal

5 Delay

# ¿Qué es un código?

- Un *alfabeto* es un conjunto de símbolos.
- Dado un alfabeto fuente  $\Sigma$ , un *código* es una correspondencia entre todas las secuencias posibles de símbolos de  $\Sigma$  a secuencias de símbolos de otro alfabeto  $X$  (alfabeto código).
- Muchas veces son utilizados a los efectos de lograr una representación más eficiente de la información (i.e., para eliminar redundancia).

# Código bloque y código no singular

- Un *código bloque* es aquél que asigna cada símbolo de  $\Sigma$  a una secuencia fija de símbolos de  $X$ :

$$C : \Sigma \rightarrow X^*$$

- Ejemplo:

- $C_1 : \{s_1, s_2, s_3, s_4\} \rightarrow \{0, 1\}^*$

$$C_1(s_1) = 0$$

$$C_1(s_2) = 11$$

$$C_1(s_3) = 01$$

$$C_1(s_4) = 101$$

- Un código  $C$  se dice *no singular* si todas sus palabras son distintas (i.e., si  $C$  es una función inyectiva).
  - $C_1$  es no singular.

# Código instantáneo y código unívocamente decodificable

- Un código es *unívocamente decodificable* si ninguna tira de símbolos del alfabeto código admite más de una única decodificación.
  - Definición más formal: si su extensión de orden  $n$  es no singular  $\forall n \in \mathbb{N}$ .
- Un código es *instantáneo* cuando es posible decodificar las palabras sin necesidad de conocer los símbolos que la suceden.
  - Condición necesaria y suficiente: ser *libre de prefijos*, no codificar ningún símbolo como prefijo de otro.

# Código instantáneo y código unívocamente decodificable

- Un código es *unívocamente decodificable* si ninguna tira de símbolos del alfabeto código admite más de una única decodificación.
  - Definición más formal: si su extensión de orden  $n$  es no singular  $\forall n \in \mathbb{N}$ .
- Un código es *instantáneo* cuando es posible decodificar las palabras sin necesidad de conocer los símbolos que la suceden.
  - Condición necesaria y suficiente: ser *libre de prefijos*, no codificar ningún símbolo como prefijo de otro.

$C_1$  no es unívocamente decodificable:



# Código instantáneo y código unívocamente decodificable

- Un código es *unívocamente decodificable* si ninguna tira de símbolos del alfabeto código admite más de una única decodificación.
  - Definición más formal: si su extensión de orden  $n$  es no singular  $\forall n \in \mathbb{N}$ .
- Un código es *instantáneo* cuando es posible decodificar las palabras sin necesidad de conocer los símbolos que la suceden.
  - Condición necesaria y suficiente: ser *libre de prefijos*, no codificar ningún símbolo como prefijo de otro.

$C_1$  no es unívocamente decodificable:

0101

# Código instantáneo y código unívocamente decodificable

- Un código es *unívocamente decodificable* si ninguna tira de símbolos del alfabeto código admite más de una única decodificación.
  - Definición más formal: si su extensión de orden  $n$  es no singular  $\forall n \in \mathbb{N}$ .
- Un código es *instantáneo* cuando es posible decodificar las palabras sin necesidad de conocer los símbolos que la suceden.
  - Condición necesaria y suficiente: ser *libre de prefijos*, no codificar ningún símbolo como prefijo de otro.

$C_1$  no es unívocamente decodificable:

0101

01 01  
⏟ ⏟  
 $s_3$   $s_3$

# Código instantáneo y código unívocamente decodificable

- Un código es *unívocamente decodificable* si ninguna tira de símbolos del alfabeto código admite más de una única decodificación.
  - Definición más formal: si su extensión de orden  $n$  es no singular  $\forall n \in \mathbb{N}$ .
- Un código es *instantáneo* cuando es posible decodificar las palabras sin necesidad de conocer los símbolos que la suceden.
  - Condición necesaria y suficiente: ser *libre de prefijos*, no codificar ningún símbolo como prefijo de otro.

$C_1$  no es unívocamente decodificable:

0101

$\underbrace{01}_{s_3} \underbrace{01}_{s_3}$

$\underbrace{0}_{s_1} \underbrace{101}_{s_4}$

# Código instantáneo y código unívocamente decodificable

- Un código es *unívocamente decodificable* si ninguna tira de símbolos del alfabeto código admite más de una única decodificación.
  - Definición más formal: si su extensión de orden  $n$  es no singular  $\forall n \in \mathbb{N}$ .
- Un código es *instantáneo* cuando es posible decodificar las palabras sin necesidad de conocer los símbolos que la suceden.
  - Condición necesaria y suficiente: ser *libre de prefijos*, no codificar ningún símbolo como prefijo de otro.

$C_1$  no es unívocamente decodificable:

0101

$\underbrace{01}_{s_3} \underbrace{01}_{s_3}$

$\underbrace{0}_{s_1} \underbrace{101}_{s_4}$

...y tampoco es instantáneo:

$$C_1(s_3) = 01 = C_1(s_1) = 0$$

# Resultado

## Teorema

Código instantáneo  $\Rightarrow$  código unívocamente decodificable

## Longitud de código

- Dado un código  $C$  sobre una fuente  $S$ , la *longitud media* de  $C$ ,  $L(C)$ , se define como

$$L(C) = \sum_{s \in S} |C(s)| \cdot P_S(s)$$

- Un código se dice *óptimo* si no existe un código para la misma fuente con menor longitud media. En otras palabras, utiliza en promedio el menor número posible de bits para codificar un mensaje.

## Longitud de código

- Dado un código  $C$  sobre una fuente  $S$ , la *longitud media* de  $C$ ,  $L(C)$ , se define como

$$L(C) = \sum_{s \in S} |C(s)| \cdot P_S(s)$$

- Un código se dice *óptimo* si no existe un código para la misma fuente con menor longitud media. En otras palabras, utiliza en promedio el menor número posible de bits para codificar un mensaje.

### Teorema: codificación sin pérdida de información

$$H(S) \leq L(C) \cdot \log r$$

Donde  $r$  es la cantidad de símbolos del alfabeto código. Para códigos binarios,

$$H(S) \leq L(C)$$

- Todo código que satisface esto se dice que codifica *sin pérdida de información*.

# Ejercicio

Para la siguiente fuente

$$S = [P(A) = 0,4; P(B) = 0,3; P(C) = 0,2; P(D) = 0,1]$$

se proponen 3 códigos posibles:

- $C_1: A = 001; B = 01; C = 11; D = 010$
- $C_2: A = 0; B = 01; C = 011; D = 111$
- $C_3: A = 1; B = 01; C = 001; D = 0001$

- a. ¿Cuáles son instantáneos?
- b. ¿Cuáles son unívocamente decodificables?
- c. De los unívocamente decodificables, ¿Cuál es más eficiente ( $H/L$ )?
- d. De los unívocamente decodificables, ¿Alguno presenta *pérdida de información*?



## Ejercicio - Primer parcial 1C2015

Sea  $S = \{s_1, \dots, s_n\}$ ,  $n > 2$ , una fuente de información de memoria nula. Supongamos que  $s_1$  es emitido con cierta probabilidad  $p_1$  y que todo otro símbolo distinto de  $s_1$  es emitido con probabilidad  $p_{\text{all}}$  (i.e.,  $P_S(s_i) = P_S(s_j) = p_{\text{all}}$ ,  $1 < i, j \leq n$ ). Sean  $C_1$ ,  $C_2$  y  $C_3$  tres códigos binarios sobre  $S$  definidos como sigue:

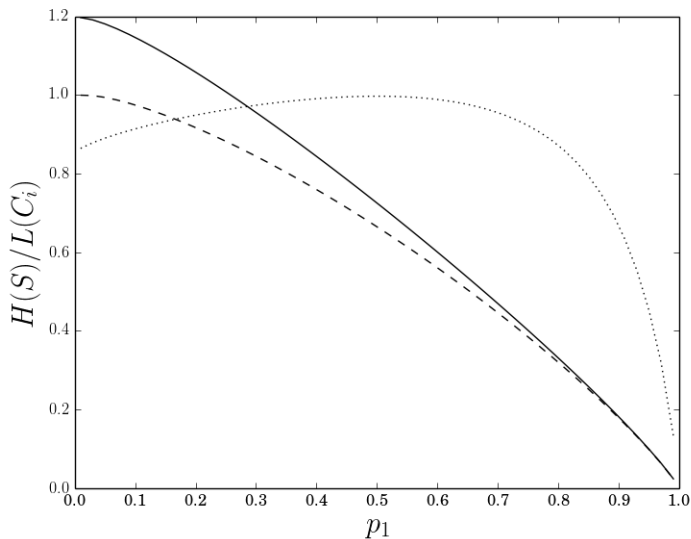
$$C_1(s_i) = \text{bin}_n(i) \quad C_2(s_i) = \begin{cases} 0 & \text{si } i = 1 \\ 1 \text{ bin}_{n-1}(i-1) & \text{si } i > 1 \end{cases} \quad C_3(s_i) = \begin{cases} \text{bin}_n(0) & \text{si } i = 1 \\ \text{bin}_{n/2}(\lfloor i/2 \rfloor) & \text{si } i > 1 \end{cases}$$

En estas definiciones,  $\text{bin}_k(j)$  indica la representación binaria de  $j$  utilizando exactamente  $\lceil \log_2(k+1) \rceil$  bits.

A partir de la siguiente gráfica de  $H(S)/L(C_i)$ , en función de  $p_1$ , para un valor fijo de  $n$ , indicar:

a. Qué curva corresponde a cada uno de los códigos.

## Ejercicio (cont.)



# Agenda

- 1 Introducción
- 2 Fuentes de información
- 3 Códigos
  - Generalidades
  - Tipos de códigos
  - Códigos sobre fuentes de memoria nula
- 4 Capacidad de canal
- 5 Delay

# Capacidad y teorema de Shannon

La *capacidad*  $C$  de un canal es la velocidad teórica máxima de transmisión, y viene dada por el...

# Capacidad y teorema de Shannon

La *capacidad*  $C$  de un canal es la velocidad teórica máxima de transmisión, y viene dada por el...

## Teorema de Shannon

$$C = B \cdot \log_2(1 + \text{SNR})$$

# Capacidad y teorema de Shannon

La *capacidad*  $C$  de un canal es la velocidad teórica máxima de transmisión, y viene dada por el...

## Teorema de Shannon

$$C = B \cdot \log_2(1 + \text{SNR})$$

- $B$ : ancho de banda (medido en Hz)

# Capacidad y teorema de Shannon

La *capacidad*  $C$  de un canal es la velocidad teórica máxima de transmisión, y viene dada por el...

## Teorema de Shannon

$$C = B \cdot \log_2(1 + \text{SNR})$$

- $B$ : ancho de banda (medido en Hz)
- SNR: relación señal-ruido (medida en *veces*)

# Capacidad y teorema de Shannon

La *capacidad*  $C$  de un canal es la velocidad teórica máxima de transmisión, y viene dada por el...

## Teorema de Shannon

$$C = B \cdot \log_2(1 + \text{SNR})$$

- $B$ : ancho de banda (medido en Hz)
- SNR: relación señal-ruido (medida en *veces*)
  - También expresable en *decibeles*. Conversión:  $\text{SNR} = 10^{\text{dB}/10}$



## Ejercicio

Considerar una señal de video en escala de grises que transmite imágenes a una resolución de  $640 \times 480$  píxeles, de los cuales cada uno puede asumir 10 niveles diferentes de brillo. Supongamos que la tasa de transmisión es de 30 imágenes por segundo y que la relación señal a ruido es de 30 dB.

- a. Calcular la entropía de la fuente si todas las imágenes fueran equiprobables.
- b. ¿Cuántos bits son necesarios para codificar cada imagen de manera óptima con un código que asigne el mismo largo a todas las imágenes?
- c. Calcular el ancho de banda mínimo requerido para soportar la transmisión de la señal resultante.

# Agenda

- 1 Introducción
- 2 Fuentes de información
- 3 Códigos
  - Generalidades
  - Tipos de códigos
  - Códigos sobre fuentes de memoria nula
- 4 Capacidad de canal
- 5 Delay

# Delay, propagación y transmisión

El *delay* representa el tiempo total que tardamos en enviar información de un punto a otro:

# Delay, propagación y transmisión

El *delay* representa el tiempo total que tardamos en enviar información de un punto a otro:

$$\text{Delay} = T_{\text{tx}} + T_{\text{prop}} + T_{\text{queue}}$$

# Delay, propagación y transmisión

El *delay* representa el tiempo total que tardamos en enviar información de un punto a otro:

$$\text{Delay} = T_{\text{tx}} + T_{\text{prop}} + \cancel{T_{\text{queue}}}$$

# Delay, propagación y transmisión

El *delay* representa el tiempo total que tardamos en enviar información de un punto a otro:

$$\text{Delay} = T_{\text{tx}} + T_{\text{prop}}$$

# Delay, propagación y transmisión

El *delay* representa el tiempo total que tardamos en enviar información de un punto a otro:

$$\text{Delay} = T_{\text{tx}} + T_{\text{prop}}$$

- $T_{\text{tx}}$ : tiempo de transmisión

# Delay, propagación y transmisión

El *delay* representa el tiempo total que tardamos en enviar información de un punto a otro:

$$\text{Delay} = T_{\text{tx}} + T_{\text{prop}}$$

- $T_{\text{tx}}$ : tiempo de transmisión
  - $= |\text{datos}| / V_{\text{tx}}$



# Delay, propagación y transmisión

El *delay* representa el tiempo total que tardamos en enviar información de un punto a otro:

$$\text{Delay} = T_{\text{tx}} + T_{\text{prop}}$$

- $T_{\text{tx}}$ : tiempo de transmisión
  - $= |\text{datos}| / V_{\text{tx}}$
- $T_{\text{prop}}$ : tiempo de propagación

# Delay, propagación y transmisión

El *delay* representa el tiempo total que tardamos en enviar información de un punto a otro:

$$\text{Delay} = T_{\text{tx}} + T_{\text{prop}}$$

- $T_{\text{tx}}$ : tiempo de transmisión
  - $= |\text{datos}| / V_{\text{tx}}$
- $T_{\text{prop}}$ : tiempo de propagación
  - $= D / V_{\text{prop}}$

# Capacidad de volumen

La *Capacidad de volumen* la cantidad de bits que entran en el medio desde que se envía el primer bit hasta que éste llega al receptor:

# Capacidad de volumen

La *Capacidad de volumen* la cantidad de bits que entran en el medio desde que se envía el primer bit hasta que éste llega al receptor:

$$C_{vol} = \text{Delay} * V_{tx}$$

## Ejercicio

Calcule la Capacidad de Volumen (cantidad de bits que entran simultáneamente) en cada uno de los siguientes medios físicos de transmisión, asumiendo que se los utiliza a su máxima Capacidad de Transmisión (es decir, sin pérdida de información):



- a.  $D = 100\text{km}$ ,  $V_{prop} = 200000\text{km/s}$ ,  $SNR = 100\text{dB}$ ,  $B = 400\text{Hz}$
- b.  $D = 100\text{km}$ ,  $V_{prop} = 200000\text{km/s}$ ,  $SNR = 10\text{dB}$ ,  $B = 400\text{kHz}$
- c.  $D = 100\text{km}$ ,  $V_{prop} = 300000\text{km/s}$ ,  $SNR = 10\text{dB}$ ,  $B = 400\text{kHz}$
- d.  $D = 100\text{m}$ ,  $V_{prop} = 300000\text{km/s}$ ,  $SNR = 10\text{dB}$ ,  $B = 400\text{kHz}$

## Ejercicio - Parcial

Un robot trepador sube escalando el monte Everest con una cámara. La señal de video de la cámara se transmite sobre un enlace inalámbrico que tiene una relación señal a ruido del orden de los 30dB, usando 50KHz del espectro. A su vez, se modela la señal de video como una fuente de información, definiendo cada símbolo como cada una de las imágenes posibles y asumiendo que son todas equiprobables.

- 1 Si queremos enviar 26 imágenes por segundo, ¿cuál es el máximo valor de entropía que puede adoptar la fuente?
- 2 ¿Cuál es la distancia para la cual el Ttx de una imagen representa el 50 % del delay? ( $V_{\text{prop}} = 300000\text{km/s}$ )

# Referencias

-  N. Abramson,  
*Teoría de la Información y Codificación*  
5ta edición.
-  W. Stallings,  
*Data and Computer Communications*  
5ta edición. Capítulo 2: Data Transmission.