## EE303: Communication Systems

Professor A. Manikas Chair of Communications and Array Processing

Imperial College London

An Overview of Fundamentals: Information Sources

## Table of Contents

1	Introduction	
2	Classification of Information Sources	
3	Discrete Memoryless Sources (DMS)	
_		

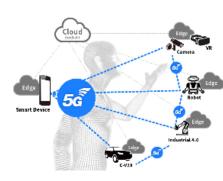
4	Measure of Information Generated by a DMS Source	10
	Source Entropy	10
	Source Information Rate	11
	<ul> <li>Examples</li> </ul>	13

5	A 'Joint' DMS S	ource	
_			

Markov Discrete Information Sources

6	Markov Discrete Information Sources	19
7	Continuous Sources/Signals	21

8	Measure of Information Generated by a Continuous Source	22
	Continuous Source: Differential Entropy     Important Relationships	22 23



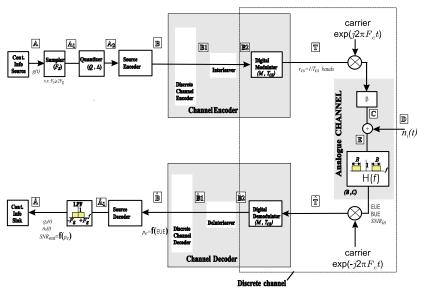
24

15

#### Introduction

- Wireline, wireless and fiber communications are all fully digital.
- The general block structure of a Digital Communication System is shown in the following page.
- it is common practice its **quality** to be expressed **in terms** of the **accuracy** with which the binary digits delivered at the output of the detector (point  $\widehat{B}2$ ) represent the binary digits that were fed into the digital modulator (point B2).
- It is generally taken that it is the fraction of the binary digits that are delivered back in error that is a measure of the quality of the communication system. This fraction, or rate, is referred to as the bit error probability, or, Bit-Error-Rate BER (point B2).

# Block Structure of a Digital Comm System



## Classification of Information Sources

- Information sources (or communication sources), or simply sources can be classified as
  - Discrete
    - ★ Discrete Memoryless Sources (MDS),
    - ★ with Memory (e.g. Markov Sources)
  - Continuous
    - ★ non-Gaussian
    - ★ Gaussian
- Examples with reference to previous page's figure:
  - ▶ continuous: up to points 🖾, 🖾, or 🗓
  - discrete:
    - ★ up to points A2 levels of quantiser, or
    - ★ up to points 🖪, 🖪 1, or 🖺 2 binary digits or binary codewords.
- The "inverse" of an information/communication source is an information/communication sink (discrete or continuous)

## Discrete Memoryless Sources

• A source is called a **Discrete Source** if produces a sequence  $\{X[n]\}$  of symbols one after another, with each symbol being drawn from a finite alphabet

$$X \triangleq \{x_1, x_2, ..., x_M\} \tag{1}$$

with a rate  $r_X$  symbols/sec,

and in which each symbol  $x_m \in X$  is produced at the output of the source with some associated probability  $\Pr(x_m)$  - abbreviated  $p_m$ , i.e.

$$\Pr(x_m) \triangleq p_m \tag{2}$$

• If successive outputs from a discrete source are **statistically independent**, or in other words, if at each instant of time the source chooses for transmission one symbol from the set  $X = \{x_1, x_2, ..., x_M\}$  and its choices are independent from one time instant to the next, then the source is called a **Discrete Memoryless Source (DMS)**.

• If  $\underline{p}$  represents the vector with elements the probabilities associated with the symbols of a source i.e.

$$\underline{p} \triangleq [p_1, p_2, ..., p_M]^T \tag{3}$$

then the set

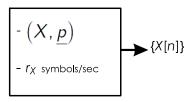
$$(X, \underline{p}) \triangleq \{(x_1, p_1), (x_2, p_2), ..., (x_M, p_M)\}$$
 (4)

with

$$\sum_{m=1}^{M} p_m = 1 \tag{5}$$

is defined as the Source Ensemble

• A DMS source can be fully described by its **ensemble**  $(X, \underline{p})$  and its **symbol rate** (symbols per second)  $r_X$ 



 The point A2 may be considered as the input of a Digital Communication System where messages consist of sequences of "symbols" selected from an alphabet e.g. levels of a quantizer or telegraph letters, numbers and punctuations.

# Measure of Information Generated by a DMS Source Source Entropy

Consider a discrete memoryless information source

$$(X, \underline{p}) = \begin{cases} (x_1, p_1) \\ (x_2, p_2) \\ \dots \\ (x_M, p_M) \end{cases} \text{ with } \sum_{m=1}^M p_m = 1$$
 (6)

Then, the average information per symbol generated by the source is given by the so-called **entropy** of the source i.e.

$$H_X \triangleq H_X(\underline{p}) = -\sum_{m=1}^{M} p_m \underbrace{\log_2(p_m)}_{\triangleq -\mathrm{I}(x_m)} \text{ bits/symbol}$$
 (7)

$$= -\underline{\rho}^{T} \log_{2}(\underline{\rho}) \quad \text{bits/symbol} \quad (8)$$

v.17.c2

#### Source Information Rate

- H<sub>X</sub> is a measure of the a priori uncertainty associated with the source output or, equivalently, a measure of the information obtained when the source output is observed.
- The notion of entropy is not restricted to the case where the ensemble is finite
   i.e. M may be ∞.
- It can also be shown that  $H_X$  is the minimum number of binary digits bits needed to encode the source output.
- Based on entropy, the average information bit-rate at the output of the source is defined as follows:

$$r_{inf} = r_X \cdot H_X \quad \text{bits/sec}$$
 (9)

- (ロ)(部)(注)(注) (注) (注) (1)(1)(1)

- Thus, we have two types of 'bits' and, therefore, two types of 'rates'.
   That is,
  - ▶ data rate :  $r_b$  in  $\frac{data\ bits}{sec}$  or simply  $\frac{bits}{sec}$
  - ▶ info rate :  $r_{inf}$  in  $\frac{information\ bits}{sec}$  or simply  $\frac{bits}{sec}$
- NB:
  - ► In general:

$$r_{\mathsf{inf}} \leq r_b$$
 (10)

▶ In ideal systems:

$$r_{\inf} = r_b \tag{11}$$

#### Example 1

• If  $X = \{0, 1\}$  with probablities

$$\underline{p} = \begin{bmatrix} Pr(0) \\ Pr(1) \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$
 (12)

then

data rate : 
$$r_b = r_X = 10 \frac{bits}{\text{sec}}$$

entropy : 
$$H_X = 1 \frac{bits}{symbol} = 1 \frac{info\ bit}{data\ bit}$$

info rate : 
$$r_{inf} = r_X . H_X = 10 \frac{bits}{sec}$$

i.e.

$$r_b = r_{\rm inf} = 10 \frac{bits}{\rm sec} \tag{13}$$

#### Example 2

• If  $X = \{0, 1\}$  with probablities

$$\underline{p} = \begin{bmatrix} Pr(0) \\ Pr(1) \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$$
 (14)

then

data rate : 
$$r_b = r_X = 10 \frac{bits}{sec}$$

entropy : 
$$H_X = 0.8813 \frac{bits}{symbol} = 0.8813 \frac{info\ bit}{data\ bit}$$

info rate : 
$$r_{inf} = r_X.H_X = 8.813 \frac{bits}{sec}$$

i.e.

$$r_b > r_{\rm inf} \tag{15}$$

### A 'Joint' DMS Source

• Consider two information sources with the ensembles  $(X, \underline{p})$  and  $(Y, \underline{q})$ , respectively, defined as follows

$$(X, \underline{p}) = \left\{ \begin{array}{c} (x_1, p_1) \\ (x_2, p_2) \\ \dots \\ (x_M, p_M) \end{array} \right\} \quad \text{with} \quad \sum_{m=1}^M p_m = 1 \qquad (16)$$

$$(Y,\underline{q}) = \begin{cases} (y_1,q_1) \\ (y_2,q_2) \\ \dots \\ (y_K,q_K) \end{cases} \text{ with } \sum_{k=1}^K p_k = 1$$
 (17)

where

$$\underline{p} = [\overline{\mathsf{Pr}(x_1)}, \overline{\mathsf{Pr}(x_2)}, ..., \overline{\mathsf{Pr}(x_M)}]^T$$
 (18)

$$\underline{q} = [\widetilde{\mathsf{Pr}(y_1)}, \widetilde{\mathsf{Pr}(y_2)}, ..., \widetilde{\mathsf{Pr}(y_K)}]^T$$
 (19)

v.17.c2

 Let us form a joined source where its symbols are taken to be pairs of symbols drawn from the two original sources X and Y.
 The new source with alphabet

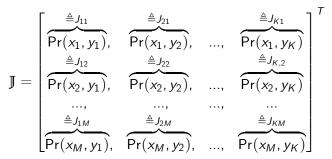
$$\{(x_1, y_1), (x_1, y_2), ..., (x_2, y_1), ..., (x_m, y_k), ..., (x_M, y_K)\}$$

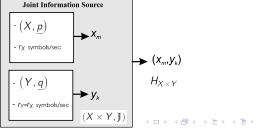
is known as **joint source** and its ensemble as **joint ensemble** defined as follows:

$$(X \times Y, \mathbb{J}) = \begin{cases} (x_{1}, y_{1}), \Pr(x_{1}, y_{1}) \\ (x_{1}, y_{2}), \Pr(x_{1}, y_{2}) \\ \dots \\ (x_{m}, y_{k}), \Pr(x_{m}, y_{k}) \\ \dots \\ (x_{M}, y_{K}), \Pr(x_{M}, y_{K}) \end{cases}$$

$$= \left\{ \left( (x_{m}, y_{k}), \underbrace{\Pr(x_{m}, y_{k})}_{=J_{km}} \right), \forall mk : 1 \leq m \leq M, 1 \leq k \leq K \right\}$$

 Note that the joint probabilistic relationship between the symbols of X and Y, described by the so-called joint-probability matrix ,





$$H_X \triangleq -\sum_{m=1}^{M} p_m \underbrace{\log_2(p_m)}_{\triangleq -\mathrm{I}(\mathsf{x}_m)} = -\underline{p}^T \log_2(\underline{p}) \qquad \frac{bits}{symbol}$$
 (21)

$$H_Y \triangleq -\sum_{k=1}^K q_k \underbrace{\log_2(q_k)}_{\triangleq -\mathrm{I}(x_m)} = -\underline{q}^T \log_2(\underline{q}) \qquad \frac{bits}{symbol}$$
 (22)

$$H_{X\times Y} \triangleq -\sum_{m=1}^{M} \sum_{k=1}^{K} J_{km} \underbrace{\log_2(J_{km})}_{\triangleq -I(x_m, y_k)}$$
(23)

$$= -\underline{1}_{K}^{T} \left( \underbrace{\mathbb{J} \odot \log_{2}(\mathbb{J})}_{K \times M \ matrix} \right) \underline{1}_{M} \ \frac{bits}{symbol}$$
 (24)

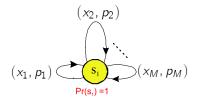
where

 $1_M = a$  column vector of M ones

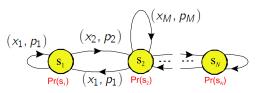
# Markov Discrete Information Sources

#### Entropy

A DMS can be modelled as with a single-state transition diagram



 Markov sources are modelled by Markov state transition diagrams of N states (N > 1), e.g.



The entropy of a Markov Source is given as follows:

$$H = \sum_{i=1}^{N} \Pr(s_i).H_i \quad \frac{bits}{symbol}$$
 (25)

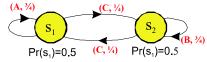
where

N = number of states

 $Pr(s_i) = Probability to be on the i<sup>th</sup> state s<sub>i</sub>$ 

 $H_i$  = entropy of the  $i^{th}$  state (considered as a DMS)

 Example: Consider a two-state Markov source which gives the symbols A,B,C according to the following model:



Then (for you)

► 
$$H_1 = ?$$

► 
$$H_2 = ?$$

H = ? bits/symbol



# Continuous Sources/Signals

- A source whose output is a continuous signal in both amplitude and time is called a Continuous Source.
- However, this definition may be relaxed to include also signals that are continuous in amplitude but discrete in time, e.g. point A2. That is the condition is relaxed to only "continuous in amplitude".
- A continuous source which relates directly to analogue signal waveforms is described by its **ensemble**  $(g(t), pdf_g(g))$  where  $pdf_g(g)$  denotes the amplitude probability density function of g(t).
- However, in order to describe fully a continuous source g(t), in addition to the source ensemble, the following parameters of g(t) should be specified/identified
  - ▶ the power spectral density  $PSD_{g}(f)$ ,
  - ▶ the bandwidth .



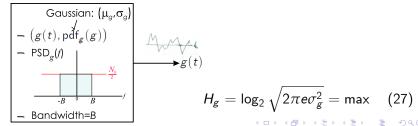
# Measure of Information Generated by a Continuous Source

#### Continuous Source: Differential Entropy

- Entropy of an analoque source  $= \infty$
- Differential entropy:

$$H_{g} \triangleq -\int_{-\infty}^{\infty} pdf_{g}(g). \log_{2}(pdf_{g}(g)).dg$$
 (26)

- The term "entropy" in this course (for analogue signals) will be used to refer to "differential entropy"
- Entropy of a Gaussian Source



#### Important Relationships

• At the output of an information source g(t) the following are very important:

Entropy: 
$$H_g$$
 (28)

$$\max(Entropy)$$
 : Gaussian Entropy =  $\log_2 \sqrt{2\pi e \sigma_g^2}$  (29)

Entropy Power : 
$$N_g \triangleq \frac{1}{2\pi e} 2^{2H_g}$$
 (30)

Average Power : 
$$P_g \triangleq \mathcal{E}\left\{g(t)^2\right\}$$
 (31)

In general: 
$$P_g \ge N_g$$
 (32)

if 
$$pdf_g = Gaussian \Rightarrow P_g = N_g$$
 (33)

if 
$$pdf_g \neq Gaussian \Rightarrow P_g > N_g$$
 (34)

## A Note on Information Sinks

- A communication sink is the destination of the symbols produced by a source and transmitted from the input to the output of a channel.
- It has one input and no output and can be seen as the inverse of a source. That is, matching its associated source, a sink is either
  - continuous, or
  - discrete
- Examples (with reference to the block structure of a comm system):
  - ightharpoonup continuous: from ponts  $\widehat{A}$ ,  $\widehat{A1}$ , or  $\widehat{T}$
  - ▶ discrete: form points B. B1 or B2
- A sink is described by
  - **1** a performance (fidelity) criterion  $\rho(X, Y)$ 
    - ★ e.g. criterion for a discrete sink: BER
    - ★ e.g. criterion for a continuous sink: SNR
  - 2 a maximum allowed distortion  $D_{\text{max}}$ (e.g.  $D_{\text{max}} = 10^{-3}$  i.e. BER $< 10^{-3}$ )

