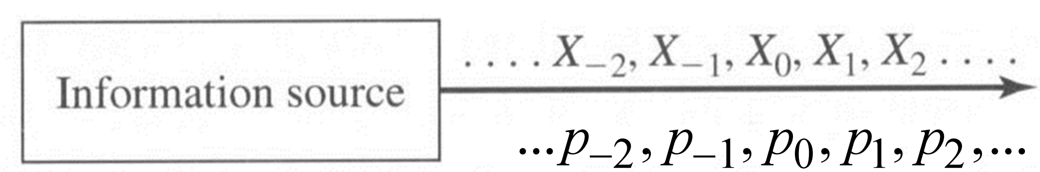
Introduction to Information Theory

- All communication systems are designed to transmit information.
- What is information? Qualitative (intuitive) description (knowledge about something) is not enough -> quantitative definition is required.
- Start with an information source, which produces outputs that are not known to the receiver in advance.
- Bandlimited analog (continuous) source -> sampling theorem -> discrete-time source.
- Simple model: discrete memoryless source (DMS):



Quantitative Measure of Information

- Defined in such a way that certain intuitive properties are satisfied.
- There is a probability associated with each source output: $[x_1 \ldots x_n]$ -> $[p_1 \ \ldots \ p_n]$. Which output conveys more information, highly probable or less probable?
- Information measure should be decreasing function of probability -> *Least probable output conveys most information*. It should also be a smooth function.

$$I_i = I(p_i), \ \ p_1 < p_2 \leftrightarrow I_1 > I_2$$

- Total information measure of two independent events is the sum of individual information measures:

$$p = p_1 p_2 \Rightarrow I(p) = I(p_1) + I(p_2)$$

Information Measure

- The only function that satisfies the properties above is the logarithm: $$I(x_i) = -\log(p(x_i)) \Rightarrow I = -\log_2 p \ [\text{bits}]$$
- The base of logarithm is not important. If the base is 2, information is measured in bits.
- Important properties:

$$I(x_i) = 0 \text{ if } p(x_i) = 1$$
$$I(x_i) \geq 0$$
$$I(x_i) > I(x_j) \text{ if } p(x_i) < p(x_j)$$
$$I(x_i x_j) = I(x_i) + I(x_j) \text{ if } x_i \ \& \ x_j \text{ are independent}$$

- Q.: What happens if p_i=0 ?

Average Information & Entropy

- Communication system: long sequences are transmitted.
- Entropy – average information content of the source per symbol:

$$H\left(X\right)=E_x\left[I(X)\right]=E_x\left[-\log p(X)\right]=-\sum_{i=1}^{N} p_i \log p_i$$

- It is a measure of uncertainty about x (on average): the more is known about x, the less is the entropy
- Example: uniform random variable,

$$x_i, i=1,\ldots,n \rightarrow p_i=1/n \rightarrow H(X)=?$$

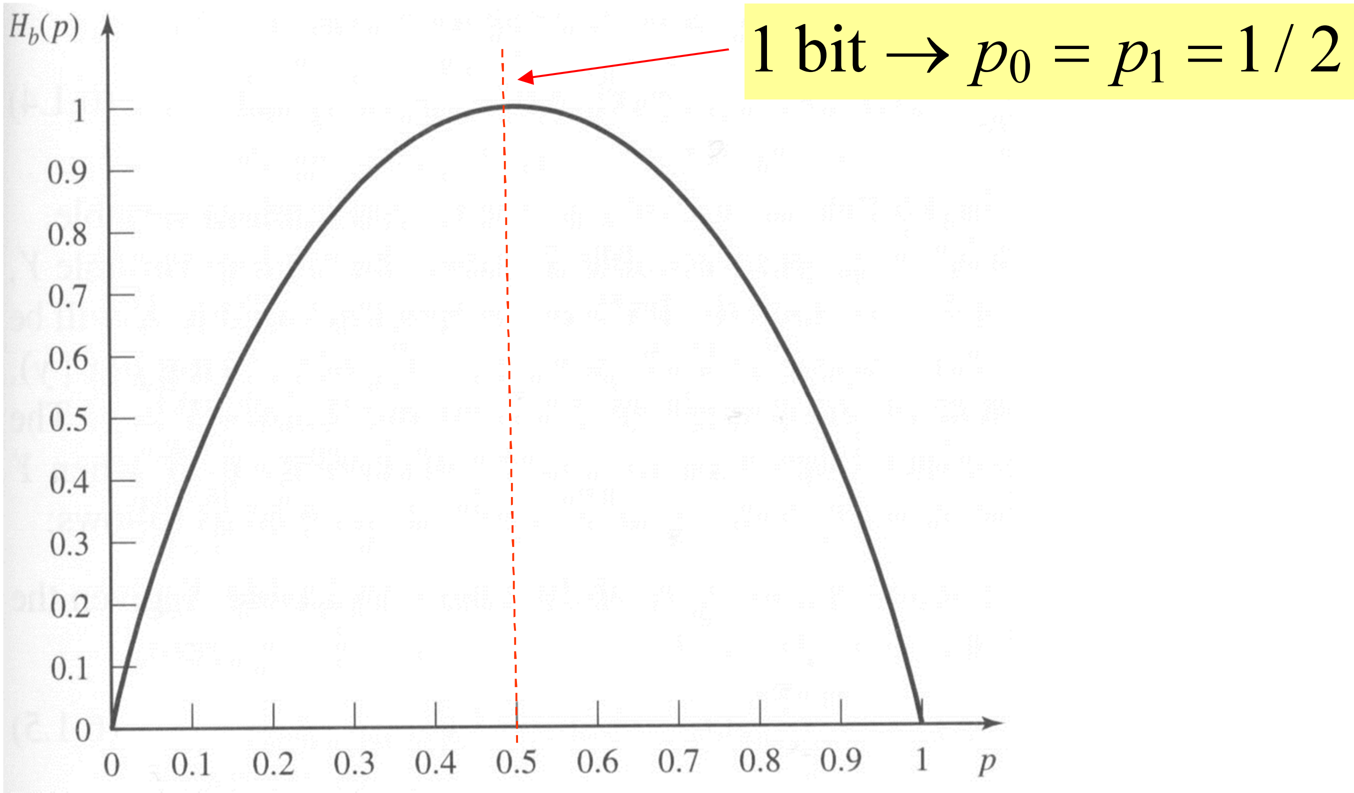
- Note that $0 \le H\left(\mathrm{X}\right) \le \log n$

When lower bound is achieved? Upper?

Example: Binary Memoryless Source

- Two possible outcomes, x_1 & x_2 -> p_1=p & p_2=1-p :

$$H(X) = -p\log p - (1-p)\log(1-p)$$



J.Proakis, M.Salehi, Communications Systems Engineering, Prentice Hall, 2002

Example: Rate of a Bandlimited Source

- Consider a baseband, bandlimited source, F_{max}=4 kHz, sampled at Nyquist rate.
- Assume that the samples are quantized to [-2,-1,0, 1, 2], and the corresponding probabilities are [1/16, 1/8, 1/2, 1/4, 1/16].
- Find the bit rate of the source. Solution:

$$H(X) = \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{8}\log 8 + \frac{2}{16}\log 16 = \frac{15}{8} \text{ bit/sample}$$

$$R = H(X)\cdot f_s = H(X)\cdot 2F_{\max} = 15 \text{ kbit/s}$$

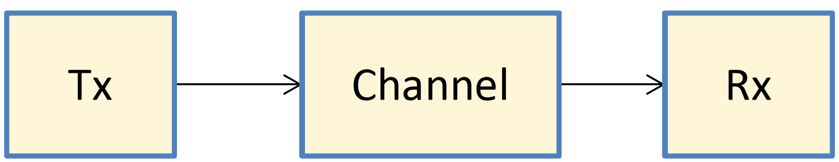
- What would be the answer for a uniformly-distributed source?

Channel Capacity

- This is the most fundamental notion in communication & information theory. It gives the fundamental limit on reliable communications over a given channel.
- Channel capacity C: maximum rate [bit/s] of reliable transmission of information over that channel.
- Reliable (error-free) transmission is possible only if

$$R_b < C$$

- The limit comes from the laws of Nature, not technological limitations.



Capacity of AWGN Channel

- Additive white Gaussian noise (AWGN) channel:

$$y = x + \xi$$

- where x – channel input, y – the output, ξ - AWG noise.
- Its capacity is (Shannon, 1948):

$$C = \Delta f \cdot \log(1 + SNR) \text{ [bit/s]}$$

- where SNR – signal-to-noise power ratio, Δf - channel bandwidth. Bandwidth can be traded for power and vise versa!
- Reliable (error-free) transmission is possible only if

$$R_b[bits/s] < C$$

- Modern systems can approach it closely.

Example: Telephone Channel

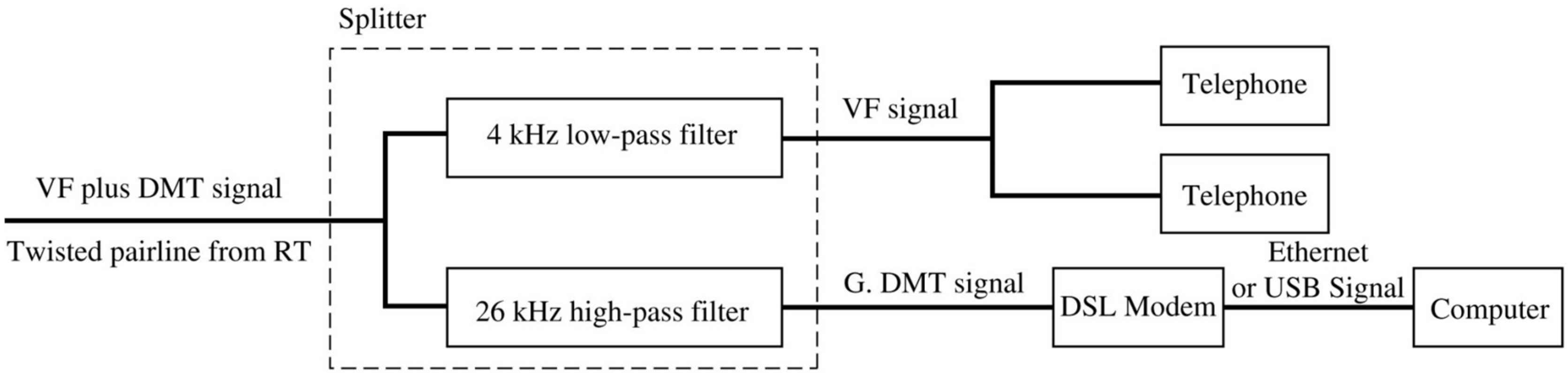
$$\Delta f = 3.4 \text{ kHz}, SNR{=}50 \text{ dB}$$

$$C = \Delta f \cdot \log(1 + SNR) \approx 56 \text{ kbit/s}$$

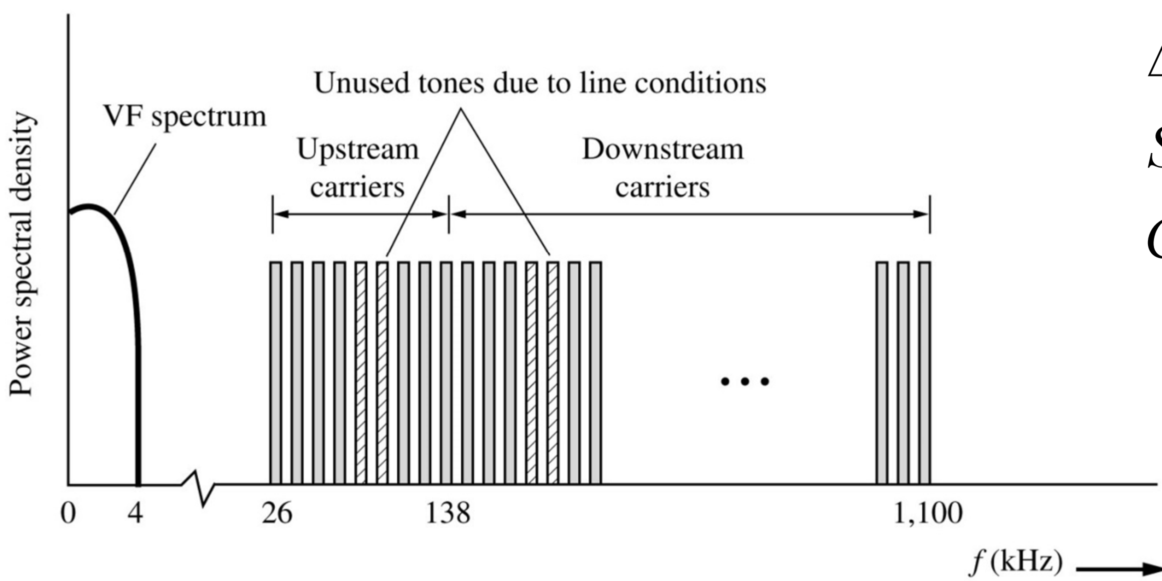
Linear SNR, not in dB!

How to increase C?

Example: Telephone Channel, DSL



(a) Customer Premises Equipment (CPE)



(b) VF plus DMT Spectrum

$\Delta f \approx 1$ MHz

SNR=50 dB

$C = ?$

Couch, Digital and Analog Communication Systems, Seventh Edition ©2007 Pearson Education

Summary

- Information source. Simple model: discrete memoryless source.
- Quantitative measure of information. Entropy.
- Rate of a discrete memoryless source.
- Communication channel model.
- Channel capacity.
- Capacity of AWGN channel.

- **Homework**: Reading, Couch, 1.9, 1.10.
- J.G. Proakis, M. Salehi, Fundamentals of Communication Systems, Perason, 2014 (or 2005), Ch. 12.1, 12.2, 12.4 – 12.6.
- Attempt to solve some examples. Follow the topics we discussed in the class.
- Extra reference: T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, 2006