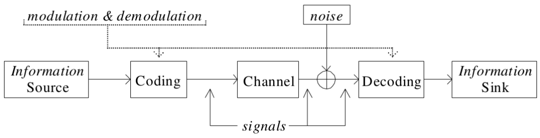
Information in Communication Systems:
Quantifying Information

Notes

Chapter Contents

Notes

Communicating Information



Information is communicated consisting of:

- **Symbols**
  - from **Symbol Set** or **Vocabulary**
- **Signals**
  - **Symbols** are coded into **Signals**

Notes

Communicating Information

But how efficiently does the system perform?

- How many bits of genuine *information* (not just binary digits) are transferred per second?
- Using what resources? e.g.
  - Channel Bandwidth
  - Power

Need to know two things:

- How many symbols are being sent per second?
- How much information does each symbol carry?

Notes

Information[1]

Now for some information theory...

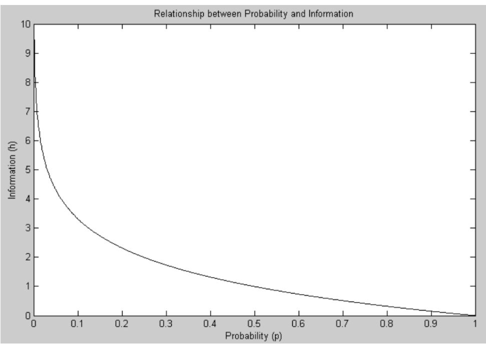
Some axioms:

1. Information is always greater than or equal to zero
2. The more unlikely a message is, the more information it carries
3. An inevitable message (i.e. $p = 1$) carries no information
4. If two messages are *independent* then their combined information content is the sum of their individual information contents.

[1]Claude Shannon developed information theory with his seminal paper from 1948: "*A mathematical theory of communication*", Bell Sys. Tech. J. vol. 27, pp. 379-423 and pp. 623-656, July and October.

Notes

Information



Notes

Information

Many functions can fulfill first 3 axioms for $h(p)$.
However all 4 axioms only satisfied by

$$h = -\log(p) = \log\left(\frac{1}{p}\right)$$

4th axiom: **If two messages are independent then their combined information content is the sum of their individual information contents.**
A logarithm is very useful for this 4th axiom because:

$$\begin{aligned} h(A,B) &= -\log(p(A,B)) = -\log(p(A)p(B)) \\ &= -\log(p(A)) - \log(p(B)) = h(A) + h(B). \end{aligned}$$

Notes

Information

What base should the logarithm take?

- Base 2 is convenient for digital systems because:
  - sample space

    $$U = \{\text{true}, \text{false}\}$$

    and assuming equal probabilities so that

    $$p(\text{true}) = 0.5 \quad \text{and} \quad p(\text{false}) = 0.5$$
  - results in

    $$h(\text{true}) = -\log_2(0.5) = 1. \quad \text{and} \quad h(\text{false}) = -\log_2(0.5) = 1.$$

So, a binary system with equal probabilities results in information of 1 bit.

Notes

Self-Information

Self information content of an outcome is measured in **bits**:

$$h(k \in U) = \log_2 \frac{1}{p_k} \quad \text{bits.}$$

We can calculate logarithm to base 2 with

$$\log_a(b) = \frac{\log_c(b)}{\log_c(a)}.$$

Many calculators and or computer math programs also offer either logarithms for any base or logarithm base 2.

Notes

Average Information

Also known as Entropy:

$$H = \sum_{k \in U} p_k h_k = -\sum_{k \in U} p_k \log_2(p_k).$$

Example
Calculate the Entropy for the following system:

$$U = \{A, B, C, D\}$$

with probabilities

Symbol	A	B	C	D
p_k	0.5	0.25	0.125	0.125

Notes

Entropy

Solution
Number of bits for each symbol $h_k = \log_2(1/p_k)$

Symbol	A	B	C	D
p_k	0.5	0.25	0.125	0.125
h_k	$1.$	$2.$	$3.$	$3.$

Product with probability $p_k h_k$

Symbol	A	B	C	D
p_k	0.5	0.25	0.125	0.125
$p_k h_k$	0.5	0.5	0.375	0.375

Entropy then:

$H = p_A h_A + p_B h_B + p_C h_C + p_D h_D = 0.5 + 0.5 + 0.375 + 0.375 = 1.75$ bits.

Notes

Max Entropy

For $U = \{0, 1\}$ (2 symbols) with a $0/1$ message set with

$$p(0) = x \quad \text{and} \quad p(1) = 1 - x$$

Entropy given by

$$H = -x \log_2(x) - (1 - x) \log_2(1 - x)$$

Need to find maximum of x which maximizes H, i.e.

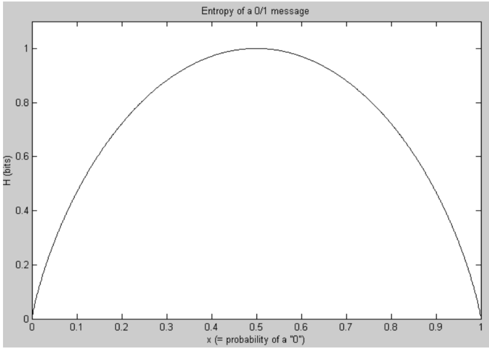
$$\max_x(H).$$

Solution given by

$$\frac{dH}{dx} = 0.$$

Notes

2-Symbol Max Entropy



Notes

2-Symbol Max Entropy

Maximum entropy occurs when $x = 0.5$, i.e.

$$p_k = 0.5 = \arg\max_x H.$$

- *Maximum* amount of information transmitted by single binary digit:

  one bit of information
- This only occurs when "0" and "1" **equally likely**, i.e.

  $p_0 = 0.5$ and $p_1 = 0.5$.
- This is true for the transmission of symbol sets of any size. i.e. Maximum entropy occurs when **all symbols equally likely**:

$$p_k = \frac{1}{N} \quad \text{and} \quad H = \log_2(N) \text{ bits.}$$

Notes

Next Time

Next time we will look at redundancy, source coding and the limiting effects of noise.

Notes